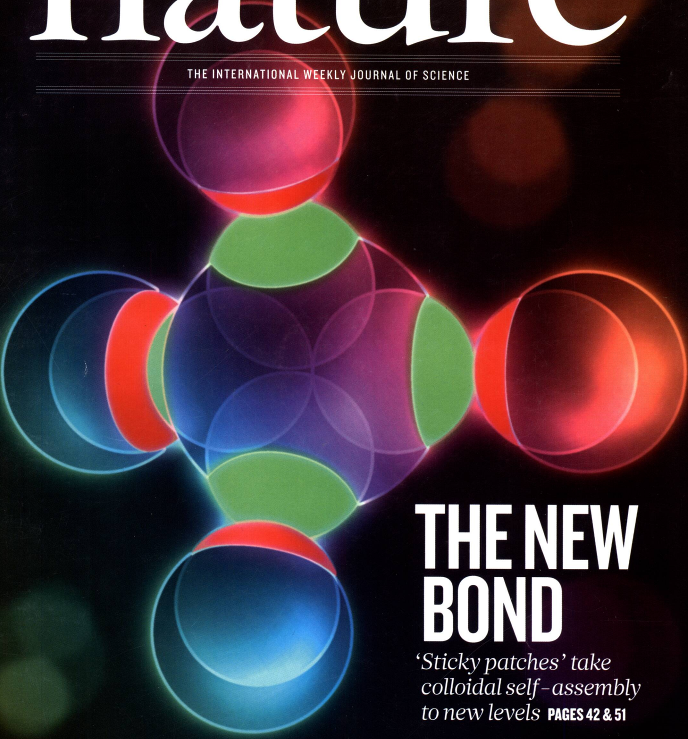
OUTLOOK
Autism

nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE



THE NEW BOND

'Sticky patches' take colloidal self-assembly to new levels **PAGES 42 & 51**

GENOMICS

EVERY CELL FOR ITSELF

Cell-by-cell sequencing reveals in-tissue diversity
PAGE 27

EXPERIMENTAL BIOLOGY

FRESH MODELS NEEDED

Research is too focused on lab-animal classics
PAGE 31

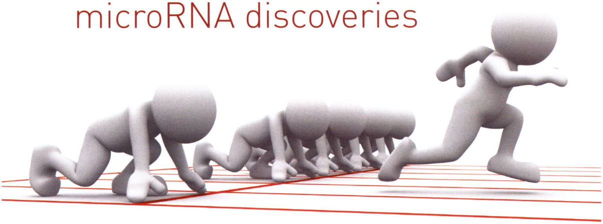
ECOLOGY

COLLATERAL DAMAGE

How common pesticides harm bumblebees
PAGES 43 & 105

NATURE.COM/NATURE
1 November 2012 £10
Vol. 491, No. 7422

Accelerate your microRNA discoveries



World leading microRNA products and services

LNA™-enhanced microRNA research tools enable reliable and accurate analysis of microRNA expression and function. Exiqon offers a complete product portfolio for your microRNA discoveries, comprehensive profiling services as well as advanced online resources:

- **Unrivaled sensitivity & specificity** – ensure accurate results for all microRNAs
- **Fast & easy workflows** – save time and increase reliability
- **Exceptional support** – get the most out of your microRNA research

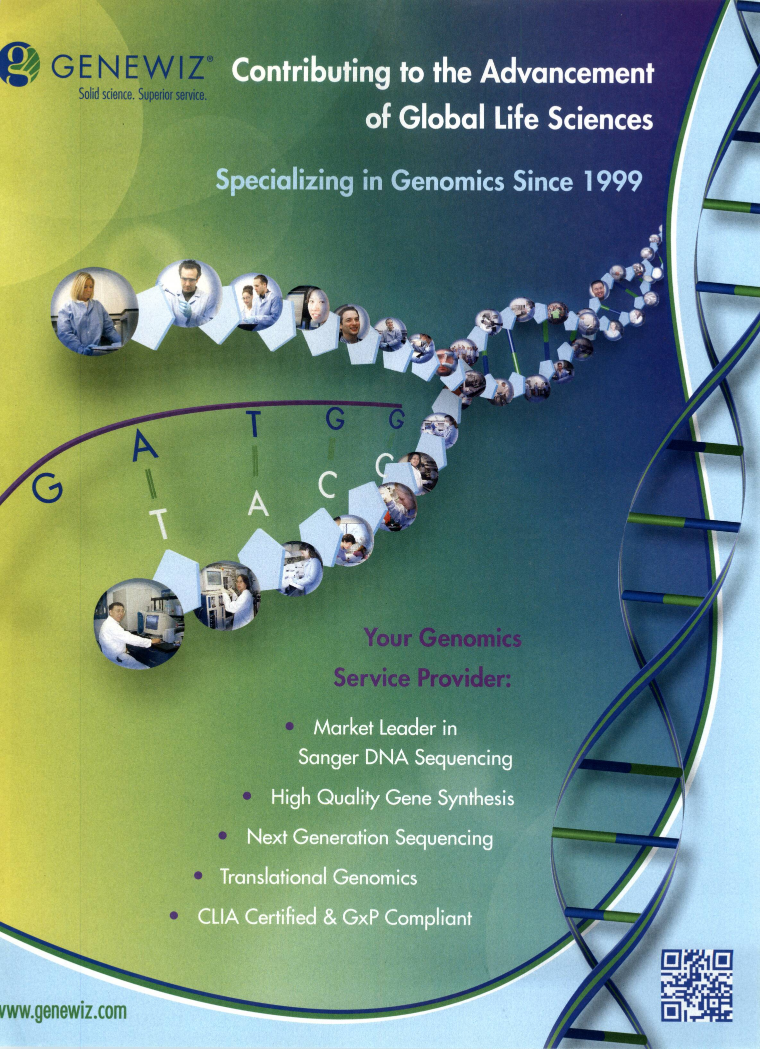
Get expert advice from Exiqon

Visit **www.exiqon.com** to learn more

Isolation	**miRCURY™ RNA Isolation Kits**
Expression Analysis	**miRCURY LNA™ microRNA Array**
Expression Analysis	**miRCURY LNA™ Univ. RT microRNA PCR**
Localization	**miRCURY LNA™ microRNA Detection**
Functional Analysis	**miRCURY LNA™ microRNA Inhibitors**

EXIQON
Seek Find Verify

NOTICE TO PURCHASER: DISCLAIMER OF LICENSE No license is conveyed with the purchase of this product under any of US Patents Nos. 5,210,015, 5,487,972, 5,804,375, 5,994,056, 6,171,785, 6,214,979, 5,538,848, 5,723,591, 5,876,930, 6,030,787, and 6,258,569, and corresponding patents outside the United States, or any other patents or patent applications, relating to the 5' Nuclease and dsDNA-Binding Dye Processes. For further information contact the Director of Licensing, Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404, USA



GENEWIZ®
Solid science. Superior service.

Contributing to the Advancement of Global Life Sciences

Specializing in Genomics Since 1999

Your Genomics Service Provider:

- Market Leader in Sanger DNA Sequencing
- High Quality Gene Synthesis
- Next Generation Sequencing
- Translational Genomics
- CLIA Certified & GxP Compliant

www.genewiz.com

eppendorf

EPPENDORF AWARD – FOR YOUNG EUROPEAN INVESTIGATORS –

Winner 2012
Elizabeth Murchison,
United Kingdom

APPLICATION DEADLINE
15 JANUARY 2013

Call for Entries!

Eppendorf Award 2013 for Young European Investigators
Since 1995 the Eppendorf Young Investigator Award is granted annually to European researchers not older than 35 years. It acknowledges outstanding contributions to biomedical research in Europe based on methods of molecular biology, including novel analytical concepts. The winner is selected by an independent expert committee chaired by Reinhard Jahn (Max Planck Institute for Biophysical Chemistry, Göttingen, Germany).

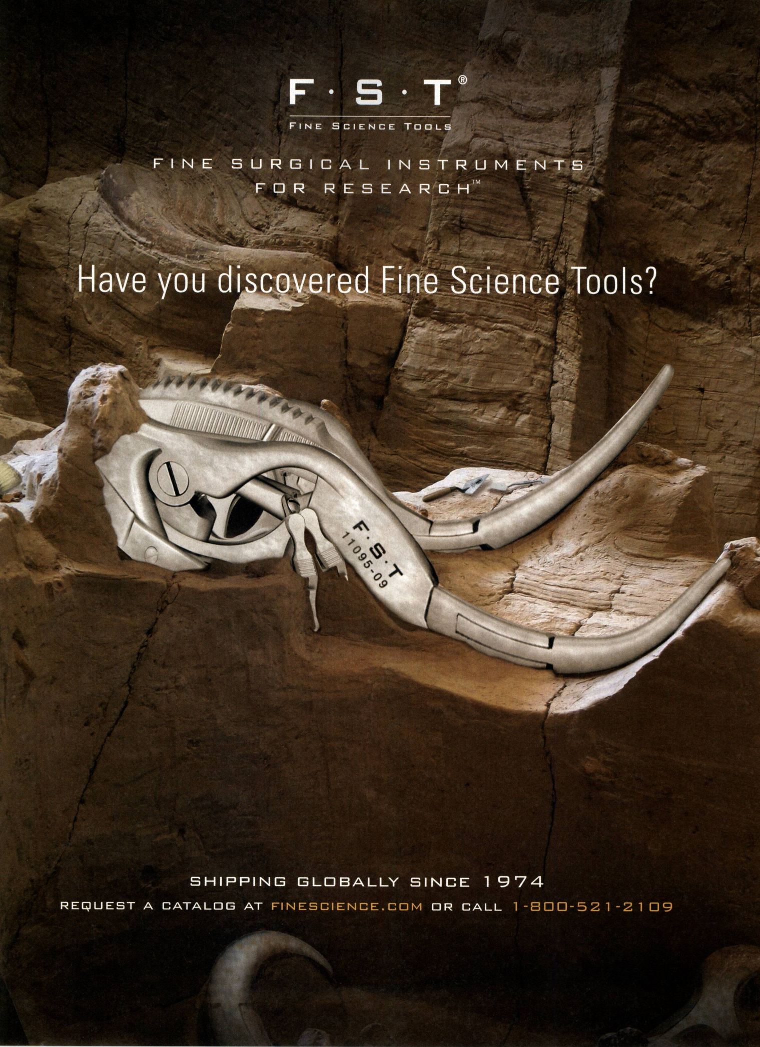
Selection criteria
1. Ability to work independently
2. Creativity
3. Successful practical application of research approach
4. Significance of research results for the field of medicine
5. Achievement of candidate in relation to his/her age

You could be next to receive
> a prize money of 15,000 Euro
> an invitation to the prize ceremony at the EMBL Advanced Training Centre in Heidelberg on 6 June 2013
> an invitation to visit Eppendorf AG in Hamburg
> coverage of your work by Nature in print and in a podcast

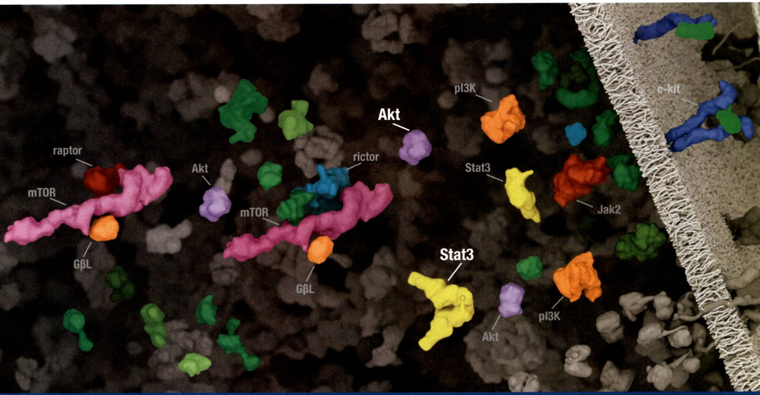
Apply online at www.eppendorf.com/award/application

Eppendorf® is a registered trademark of Eppendorf AG. All rights reserved, incl. graphics and images. Copyright ©2012 by Eppendorf AG.

The Eppendorf Award is presented in partnership with
nature


F·S·T®
FINE SCIENCE TOOLS
FINE SURGICAL INSTRUMENTS FOR RESEARCH™
Have you discovered Fine Science Tools?
F·S·T
11095-09
SHIPPING GLOBALLY SINCE 1974
REQUEST A CATALOG AT FINESCIENCE.COM OR CALL 1-800-521-2109

Antibodies and Related Reagents for Signal Transduction Research



Conjugated Antibodies for Flow Cytometry *from Cell Signaling Technology*

Unparalleled product quality, validation, and technical support

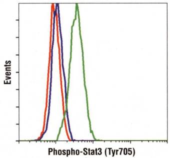
:: The highest quality antibodies provide you with the brightest signal and the lowest background.

:: Extensive validation by our in-house flow cytometry group means optimization is not left up to you, the user.

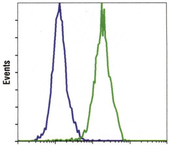
:: Technical support provided by our flow cytometry specialists translates into a thorough, fast, and accurate response.

:: High quality custom conjugations of our off-the-shelf antibodies are also available, please inquire about sizing and pricing options.

:: Multiple fluorochrome options are available to meet your unique multiplexing needs.



Phospho-Stat3 (Tyr705) (D3A7) XP® Rabbit mAb (Alexa Fluor® 488 Conjugate) #4323: Flow cytometric analysis of Jurkat cells, untreated (blue) or IFN-α treated (green), using #4323 compared to nonspecific control antibody (red).



Phospho-Akt (Ser473) (D9E) XP® Rabbit mAb (PE Conjugate) #5315: Flow cytometric analysis of Jurkat cells, untreated (green) or treated with LY294002 (PI3 Kinase Inhibitor) #9901, Wortmannin (PI3 Kinase Inhibitor) #9551, and U0126 (MEK1/2 Inhibitor) #9903 (blue), using #5315.

Alexa Fluor® is a trademark of Molecular Probes, Inc.

for quality products you can trust...

www.cellsignal.com

Cell Signaling TECHNOLOGY®

Orders (toll-free) 1-877-616-2355 | **Technical support** (toll-free) 1-877-678-8324 support@cellsignal.com | **Inquiries** info@cellsignal.com | **Environmental Commitment** eco.cellsignal.com

For Research Use Only. Not For Use In Diagnostic Procedures.

CONTENTS

1 November 2012 / Vol 491 / Issue No 7422

THIS WEEK

EDITORIALS

7 ITALY
Murky manoeuvres
Italy's politicians must adjust their attitude towards science

7 JAPAN
Bad press
Vigilance and a global perspective are key to science communication

8 SCIENCE AND SOCIETY
Fight the power
Eritrea illustrates the need for academia to maintain its independence

8 ANNOUNCEMENT
***Nature* metrics**
Online counts for research papers

WORLD VIEW

9 **How women scientists fare in the Arab world**
Rana Dajani
True equality requires recognition of women's roles in the family

RESEARCH HIGHLIGHTS

10 SELECTIONS FROM THE SCIENTIFIC LITERATURE
Resistant *C. difficile* / 'Black widow' pulsar exposed / Climate engineering boost / Do bears count in the woods?

SEVEN DAYS

12 THE NEWS IN BRIEF
China goes nuclear / US east coast hit by cyclone

CAREERS

149 FOOD SCIENCE
A smorgasbord of opportunity
There are good careers to be had in food science — and you won't go hungry

151 COLUMN
Career fast-forward
Katharina Genreith
Group leadership can be rewarding

151 CAREER BRIEFS

NATUREJOBS ADVERTISING FEATURE
Spotlight on genetics

NEWS IN FOCUS

15 SCIENTIFIC ADVICE
L'Aquila trial result sparks global backlash

16 ENERGY
Siemens' exit from solar market puts Sahara project in doubt

18 ANIMAL RESEARCH
NIH struggles to find homes for retired lab chimps

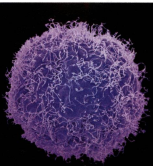
19 PHYSICS
Mathematical framework offers clues on how particles cope with gravity fluctuations

21 IMMUNOLOGY
Vaccine for hepatitis E is launched

22 FOOD SAFETY
Bid to curb acrylamide in food falters

FEATURES

24 DEVELOPING WORLD
Eritrea's shattered science
Government clampdown quashes medical success in nascent scientific enterprise



GENOMICS

Single life

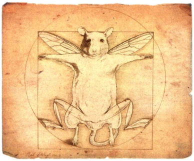
How sequencing individual cells is changing views on the body as a whole

PAGE 27

TECHNOLOGY

143 EPIGENETICS
Reading the second genomic code
DNA methylation and other epigenetic mechanisms are tempting targets for drug developers

COMMENT



31 EXPERIMENTAL BIOLOGY
There's more to life than rats and flies
Jessica Bolker
Over-reliance on familiar animal models limits progress in many areas of research

33 GENETIC TESTS
Politics and fetal diagnostics collide
Jaime S King
Non-invasive prenatal genetic tests could fall victim to anti-abortion lobby

35 POLICY
Base sustainable development goals on science
Gisbert Glaser
The UN working group needs to do its research

BOOKS & ARTS

36 NEUROSCIENCE
Found in translation
Charles Fernyhough

37 BOOKS IN BRIEF

38 HISTORY
Realms of gold
Jennifer Rampling

39 GEOGRAPHY
Going all the way
Andrew Robinson

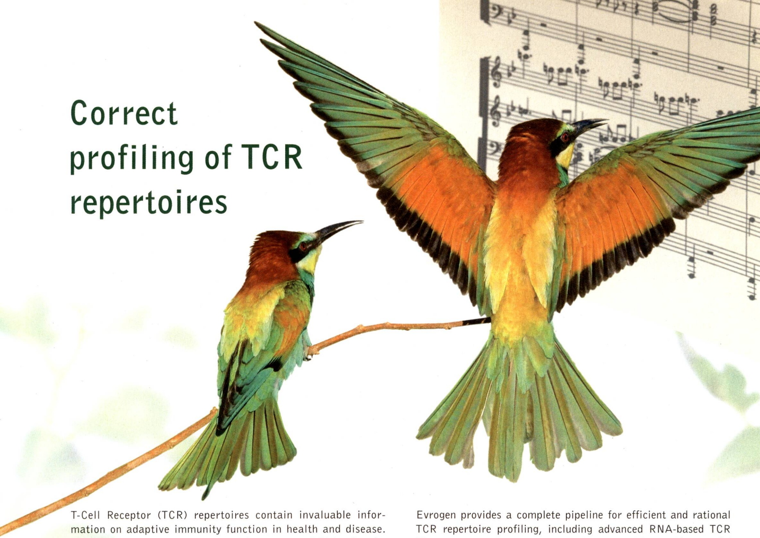
CORRESPONDENCE

40 A citable effect / Call for ZooBank registration / Ecology in Ecuador / Open collaboration on synthetic probes

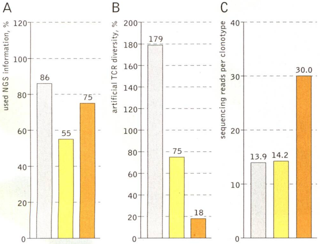
FUTURES

154 **Midnight in the cathedral of time**
Preston Grassmann

Correct profiling of TCR repertoires



T-Cell Receptor (TCR) repertoires contain invaluable information on adaptive immunity function in health and disease. Next-generation sequencing provides new capabilities for the massive analysis of TCR repertoires. However, accumulated PCR and sequencing errors, library preparation bottlenecks, and uneven PCR efficiency lead to loss of the original TCR repertoire, generation of huge artificial TCR diversity, and biased quantification of the T-cell clones. These obstacles complicate rational analysis of generated sequencing information which is crucial for accurate data interpretation in the adaptive immunity studies.



Evrogen provides a complete pipeline for efficient and rational TCR repertoire profiling, including advanced RNA-based TCR library preparation, next-generation sequencing, and powerful bioinformatic analysis with minimal loss of information originally contained in a complex lymphocyte sample. The advantages:

- Analysis of TCR beta or TCR alpha in Human or Mouse
- Minor subpopulations can be analyzed with as little as 10,000 T-cells starting material
- Deep quantitative profiling based on multiplex-free PCR and Illumina sequencing[1]
- Advanced sequencing error correction for your Illumina, 454, or Ion Torrent data[2] (Fig. 1)
- Easy-to-use analysis software with various filtering, spectratyping, and exporting options.

---

1. Mamedov *et al.* Quantitative tracking of T cell clones after haematopoietic stem cell transplantation. EMBO Mol Med. 2011 Apr; 3(4):201-7.
2. Bolotin *et al.* Next generation sequencing for TCR repertoire profiling: Platform-specific features and correction algorithms. Eur J Immunol. 2012 Jul 18. doi: 10.1002/eji.201242517.

Figure 1. Basic parameters of TCR beta repertoire analysis for Illumina data (1 million sequencing reads, human blood sample).

(A) NGS information: percentage of sequencing reads successfully mapped to final CDR3 clonotypes. (B) Artificial TCR beta variants diversity: percentage of artificial diversity generated as a result of PCR and sequencing errors over the estimated natural TCR diversity. (C) Quantitative power: average number of sequencing reads per clonotype.

Data are shown for the raw non-corrected data (gray), after filtering off the low quality sequencing reads (yellow), and after advanced correction analysis (orange).

evrogen

Evrogen JSC, Moscow, Russia
Tel.: +7(495)9884084
Fax: +7(495)9884085
e-mail: evrogen@evrogen.com
Web site: www.evrogen.com

CONTENTS

1 November 2012 / Vol 491 / Issue No 7422

RESEARCH

NEW ONLINE

41 Papers published this week at nature.com

NEWS & VIEWS

42 MATERIALS SCIENCE
Self-assembly gets new direction
Specific, directional bonding between colloid particles
Matthew R. Jones & Chad A Mirkin
SEE ARTICLE P. 51

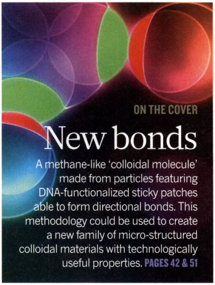
43 ECOLOGY
Bumblebees and pesticides
Chemicals have effects at both the individual-bee and colony level
Juliet L Osborne **SEE LETTER P. 105**

45 ASTEROIDS
Dark and stormy weather
Vesta's surface is contaminated with exogenous carbon-rich material
Beth Ellen Clark **SEE LETTERS PP. 79 & 83**

46 NEUROSCIENCE
Sleep to oblivion
Anaesthetics induce sleep-promoting neurons in mouse brain
Sadaf Shadan

47 OPTICS
Nanotube holograms
Two-dimensional image generated by an array of carbon nanotubes
Stéphane Larouche & David R Smith

48 ORIGINS OF LIFE
The cooperative gene
Cooperation between self-assembling RNA molecules
James Attwater & Philipp Holliger
SEE ARTICLE P. 72



ON THE COVER

New bonds

A methane-like 'colloidal molecule' made from particles featuring DNA-functionalized sticky patches able to form directional bonds. This methodology could be used to create a new family of micro-structured colloidal materials with technologically useful properties. **PAGES 42 & 51**

ASTRONOMY

Dawn raid

A detail view of protoplanet Vesta from NASA's space probe Dawn **PAGE 79**

ARTICLES

51 MATERIALS SCIENCE **Colloids with valence and specific directional bonding**
Y Wang et al. **SEE N&V P. 42**

56 GENOMICS **An integrated map of genetic variation from 1,092 human genomes**
The 1000 Genomes Project Consortium

66 STEM CELLS **Generation of functional thyroid from embryonic stem cells**
F Antonica et al.

72 ORIGINS OF LIFE **Spontaneous network formation among cooperative RNA replicators**
N Vaidya et al. **SEE N&V P. 48**

LETTERS

79 ASTRONOMY **Distinctive space weathering on Vesta from regolith mixing processes**
C M Pieters et al. **SEE N&V P. 45**

83 ASTRONOMY **Dark material on Vesta from the infall of carbonaceous volatile-rich material**
T B McCord et al. **SEE N&V P. 45**

87 PHYSICS **Observation of spatially ordered structures in a two-dimensional Rydberg gas**
P Schauß et al.

92 GEOLOGY **Fluvial response to abrupt global warming at the Palaeocene/Eocene boundary**
B Z Foreman, P L Heller & M T Clementz et al.

96 GEOPHYSICS **The elusive Hadean enriched reservoir revealed by ^{142}Nd deficits in Isua Archaean rocks**
H Rizo et al.

101 GEOPHYSICS **Fault healing promotes high-frequency earthquakes in laboratory experiments and on natural faults**
*G C McLaskey, A M Thomas, S D Glaser & R M Nadeau*et al.

105 ECOLOGY **Combined pesticide exposure severely affects individual- and colony-level traits in bees**
R J Gill, O Ramos-Rodriguez & N E Raine et al. **SEE N&V P. 43**

109 NEUROSCIENCE **Intrinsically determined cell death of developing cortical interneurons**
D G Southwell et al.

114 GENOMICS **In vivo genome editing using a high-efficiency TALEN system**
V M Bedell et al.

119 DISEASE **Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease**
A list of authors and their affiliations appears at the end of the paper

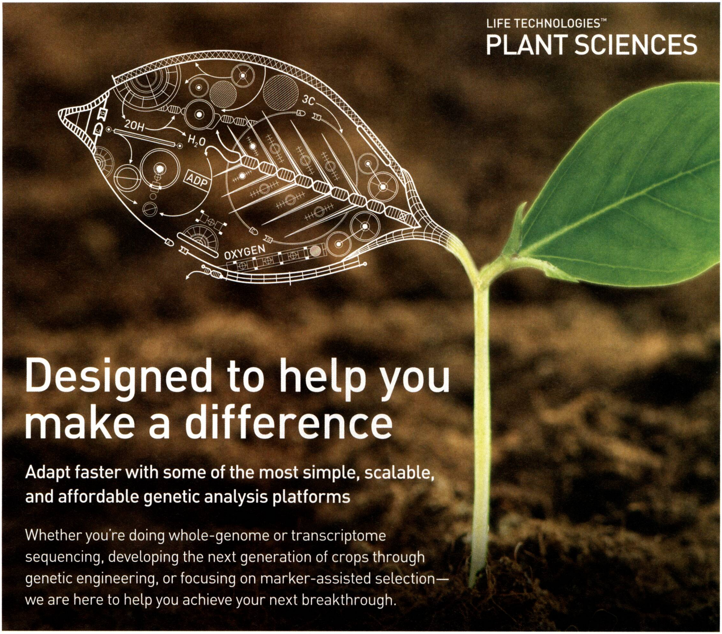
125 VIROLOGY **Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11**
M Li et al.

129 IMMUNOLOGY **Vaccine-induced CD8$^+$ T cells control AIDS virus replication**
P A Mudd et al.

134 ENVIRONMENT **The molecular basis of phosphate discrimination in arsenate-rich environments**
M Elias et al.

138 STRUCTURAL BIOLOGY **The spatial architecture of protein function and adaptation**
R N McLaughlin Jr, F J Poelwijk, A Raman, W S Gosal & R Ranganathan

PENGCHENG SONG, YU WANG & YUFENG WANG



LIFE TECHNOLOGIES™
PLANT SCIENCES

Designed to help you make a difference

Adapt faster with some of the most simple, scalable, and affordable genetic analysis platforms

Whether you're doing whole-genome or transcriptome sequencing, developing the next generation of crops through genetic engineering, or focusing on marker-assisted selection—we are here to help you achieve your next breakthrough.

Whole-genome and transcriptome sequencing

Ion Proton™ Sequencer

Marker-assisted selection and plant genotyping

QuantStudio™ 12K Flex Real-Time PCR System and Ion PGM™ Sequencer

Plant genetic engineering

GeneArt® Algae Engineering Kits

Unlock the power of plants at **lifetechnologies.com/plants**

life technologies™

For Research Use Only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of Life Technologies Corporation or their respective owners. ©2012 Life Technologies Corporation. All rights reserved. CO03022 1012

Invitrogen™ Applied Biosystems® Gibco® Molecular Probes® Novex® TaqMan® Ambion® Ion Torrent™

OUTLOOK

Supplement to 1 November 2012 / Vol 491 / Issue No 7422

AUTISM

natureOUTLOOK

AUTISM

Produced with support from:
Nancy Lurie Marks Family Foundation,
the Simons Foundation, Roche, Autism Speaks,
and The Autism Science Foundation

Illuminating a diverse disorder

Turn to page 40

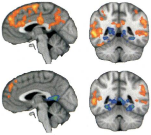
S2 EPIDEMIOLOGY
Complex disorder
Awareness, funding and research into autism are on the rise

S4 GENETICS
Searching for answers
Multidisciplinary collaborations aim to solve autism's complexity

S7 CHILD DEVELOPMENT
The first steps
Studying the siblings of children with autism points to better detection

S10 ADULTHOOD
Life lessons
Research is maturing along with its participants

S12 DIAGNOSIS
Redefining autism
Draft diagnostic guidelines are a matter of contention



Tracking the development of children is helping to define autism and the factors at play (page S7).

S14 TREATMENTS
In the waiting room
Clinicians are eager for new drugs developed specially to treat the symptoms of autism

S17 PERSPECTIVE
Imaging autism
Nicholas Lange cautions against using brain scans to diagnose autism before the underlying biology is understood

S18 CULTURE
Diverse diagnostics
Assessing a child's behaviour to diagnose autism can be done only in the context of cultural norms

S20 PERSPECTIVE
Brain scans need a rethink
Evidence for a leading theory on the causes of autism is flawed, say Ben Deen and Kevin Pelphrey

MORE ONLINE

NATURE.COM

- The Autism Enigma, *Nature* and several sister journals publish news and comment on the efforts to understand autism: go.nature.com/ypo9f8

COLLECTION

- Visit our online edition, which includes an exclusive collection of research articles on autism: go.nature.com/suc7L5

nature

1 November 2012 / Vol 491 / Issue No. 7422

EDITORIAL

LONDON, MUNICH, WASHINGTON DC, NEW YORK, BOSTON, PARIS, SAN FRANCISCO, TOKYO **nature@nature.com**

EDITOR-IN-CHIEF: Philip Campbell
EXECUTIVE EDITOR: Nick Campbell
PUBLISHING EXECUTIVE EDITOR: Maxine Clarke **EXECUTIVE EDITOR:** Véronique Kiermer
CHIEF MAGAZINE EDITOR: Tim Appenzeller
THIS WEEK/NEWS/FEATURES: Alison Abbott, David Adam, Ananyo Bhattacharya, Geoff Brumfiel, Declan Butler, Ewen Callaway, Erika Check Hayden, Daniel Cressey, David Cyranoski, Natasha Gilbert, Eric Hand, Heidi Ledford, Corie Lok, Brendan Maher, Emma Marris, Richard Monastersky, Brian Owens, Helen Pearson, Mark Peplow, Eugenie Samuel Reich, Ivan Semeniuk, Quirin Schiermeier, Jeff Tollefson, Richard Van Noorden, Meredith Wadman, M Mitchell Waldrop **COMMENT:** Sara Abdulla, Joanne Baker, Rosalind Cotter, Barbara Kiser, Alison McCook, Lucy Odling-Smee, Sarah Tomlin **NATURE PODCAST/VIDEO:** Adam Rutherford, Thea Cunningham, Kerri Smith, Charlotte Stoddart **NEWS AND VIEWS:** Sadaf Shadan, Ana Lopes, Andrew Mitchinson, Cesar Sanchez, Marian Turner **RESEARCH:** Ritu Dhand, Karl Ziemelis, Francesca Cesari, I-han Chou, Tanguy Chouard, Rosamund Daw, Alex Eccleston, Angela Eggleston, Joshua Finkelstein, Henry Gee, Patrick Goymer, Noah Gray, Marie-Thérèse Heemels, Magdalena Helmer, Karen Howell, Claudia Lupp, Barbara Marte, Andrew Mitchinson, Juliane Mössinger, Deepa Nath, Wayne Peng, Leslie Sage, Magdalena Skipper, Clare Thomas, Maria Trajkovska, John VanDecar, Liesbeth Venema, Ursula Weiss, Michael White **INSIGHTS/REVIEWS/PERSPECTIVES:** Ursula Weiss **CAREERS:** Gene Russo, Karen Kaplan **TECHNOLOGY FEATURES:** Vivien Marx **FUTURES:** Colin Sullivan **SUPPLEMENTS:** Herb Brody, Michelle Grayson, Tony Scully **SUBEDITORS:** Colin Sullivan, Mary Abraham, Sarah Archibald, Nicola Bailey, Nicolas Fanget, Isobel Flanagan, Paul Fletcher, Anne Haggart, Sasha Lewis, Dinah Loon, Louisa Lyon, Anna Novitzky, Yukie Ozawa, Katherine Perry, David Price, Jenny Rooke, David Winters, Anna York **PRODUCTION:** Jenny Henderson, Leonora Dawson-Bowling, Dan Foley, Stewart Fraser, Susan Gray, Simon Gribbin, Hazel Mayhew, James McQuat, Emilia Orviss, Yvonne Strong, Robert Sullivan, Roberta Tarran, Charles Wenz **ART AND DESIGN:** Kelly Buckheit Krause, Wesley Fernandes, Madeline Hutchinson, Barbara Izdebska, Paul Jackman, Jasiek Krzysztofiak, Annthea Lewis, Chris Ryan, Nik Spencer, Claire Welsh, Mark Wood **JAPAN EDITION:** Eiji Matsuda, Mika Ishida, Mitsuyo Utsugi, Sayaka Ayame **PRESS OFFICE:** Ruth Francis, Neda Afsarmanesh, Lisa Boucher, Jenny Chapman, Rachel Twinn, Rebecca Walton **ADMINISTRATION:** Paul Sng, Liz Bankes, Roseann Campbell, Peter Cary, Ffion Cleverley, Dana Coles, Faye Fornasier, Ashlee Gagui, Peter Hewitt, Caroline McLean, Anastasia Panoutsou, Kenneth Simpson, Michelle Viscuso, Diane Yorke

PUBLISHING

LONDON **feedback@nature.com**

MANAGING DIRECTOR & PUBLISHER: Steven Inchcoombe
ASSOCIATE PUBLISHER: Arash Hejazi
PUBLISHING PROJECT MANAGERS: Claudia Deasy, Christian Manco

TOKYO **feedback@natureasia.com**

REGIONAL MANAGING DIRECTOR — ASIA-PACIFIC: David Swinbanks
DIRECTOR — ASIA-PACIFIC: Antoine E Bocquet
OPERATIONS DIRECTOR — ASIA-PACIFIC: Hiroshi Minemura

DISPLAY ADVERTISING, SPONSORSHIP & NATUREJOBS

GLOBAL HEAD: Andrew Douglas a.douglas@nature.com

DISPLAY ADVERTISING AND SPONSORSHIP

GLOBAL HEAD: Gerard Preston g.preston@nature.com
ASSISTANT SPONSORSHIP MANAGER: Reya Silao Tel: +44 (0)20 7843 4977 (r.silao@nature.com)

EUROPE/REST OF WORLD display@nature.com Tel: +44 (0)20 7843 4960

GERMANY/SWITZERLAND/AUSTRIA: Sabine Hugi-Fürst Tel: +41 52761 3386
UK/IRELAND/SCANDINAVIA/SPAIN/PORTUGAL: Evelina Rubio Hakansson Tel: +44 (0)20 7843 4079
FRANCE/BELGIUM/THE NETHERLANDS/LUXEMBOURG/ITALY/OTHER EUROPE: Nicola Wright Tel: +44 (0)20 7843 4959

NORTH AMERICA **display@natureny.com**

NEW ENGLAND/ EASTERN CANADA: Sheila Reardon Tel: +1 617 494 4900
NEW YORK/MID-ATLANTIC/SOUTHEAST: Jim Breault Tel: +1 212 726 9334
MIDWEST: Mike Rossi Tel: +1 212 726 9255
WEST COAST/ WESTERN CANADA: George Lui Tel: +1 415 781 3804
WEST COAST NORTH: Kayla McCutchan Tel: +1 415 403 9038

ASIA-PACIFIC **display@natureasia.com**

ASIA-PACIFIC HEAD: Kate Yoneyama
JAPAN: Ken Mikami Tel: +81 3 3267 8765
GREATER CHINA/SINGAPORE: Gloria To Tel: +852 2811 7191
INDIA: Vikas Chawla Tel: +91 124 288 1057

SPONSORSHIP

BUSINESS DEVELOPMENT MANAGER: David Bagshaw Tel: +1 212 726 9215
BUSINESS DEVELOPMENT MANAGER: Patrick Murphy Tel: +1 617 475 9216

NATUREJOBS

naturejobs@nature.com

EUROPEAN HEAD OFFICE, LONDON: Tel: +44 (0)20 7843 4961
EUROPEAN SALES MANAGER: Nils Moeller (n.moeller@nature.com) **BUSINESS DEVELOPMENT MANAGER, MIDDLE EAST:** Jon Giuliani (j.giuliani@nature.com) **ADVERTISING PRODUCTION MANAGER:** Stephen Russell **NATUREJOBS ONLINE PRODUCTION:** Dennis Chu
US HEAD OFFICE, NEW YORK: naturejobs@natureny.com **US SALES MANAGER:** Ken Finnegan
JAPAN HEAD OFFICE, TOKYO: Tel: +81 3 3267 8765 **ASIA-PACIFIC HEAD:** Kate Yoneyama
ASIA-PACIFIC SALES & BUSINESS DEVELOPMENT MANAGER: Yuki Fujiwara **BUSINESS DEVELOPMENT MANAGER, GREATER CHINA, SINGAPORE:** Gloria To Tel: +852 2811 7191
INDIA: Vikas Chawla Tel: +91 1242881057 (v.chawla@nature.com)

MARKETING & SUBSCRIPTIONS

USA/CANADA/LATIN AMERICA **subscriptions@us.nature.com**
Tel: (USA/Canada)+1 866 363 7860; (outside USA/Canada) +1 212 726 9223
MARKETING: Sara Girard **FULFILMENT:** Karen Dawson

JAPAN/CHINA/KOREA **subscriptions@natureasia.com**
Tel: +81 3 3267 8751 **MARKETING:** Sachiyo Ikeda **FULFILMENT:** Masayo Tanaka

EUROPE/REST OF WORLD **subscriptions@nature.com**
Tel: +44 (0)1256 329242
MARKETING: Elena Woodstock, Marcus Parker, Hannah Phipps **FULFILMENT:** Jean Burrows

INDIA **npgindia@nature.com**
Tel: +91 124 2881053/54
HEAD OF BUSINESS DEVELOPMENT, INDIA: Debashish Brahmachari **MARKETING:** Harpal Singh Gill

ANNUAL SUBSCRIPTIONS (UK/Europe, including post and packing)
FULL RATE: £1,400
PERSONAL RATES: UK £135; Europe €209; Rest of World £210
STUDENT RATES: UK £79; Europe €122; Rest of World £125
Individual rates apply to payments by personal cheque or credit card. Back issues: £20.00.

ANNUAL SUBSCRIPTIONS (USA, including post and packing)
INSTITUTIONAL/CORPORATE RATE: $2,730 ; **PERSONAL RATE:** $199; **STUDENT RATE:** $99; **POSTDOC RATE:** $119
Printed in USA. Individual rates available only to subscribers paying by personal check or credit card. Orders for student/postdoc subscriptions must be accompanied by a copy of student ID. Rates apply to USA, Canada, Mexico/Central & South America. Add 7% GST tax in Canada (Canadian GST number 140911595). **BACK ISSUES:** US$20.00.

ANNUAL SUBSCRIPTIONS (Asia, including post and packing)
JAPAN: **年間購読料金：** *Nature のみ* 437,800 円 +21,890 円（消費税）=459,690 円（合計）
Nature + Nature Digest 466,550 円 +23,328 円（消費税）=489,878（合計）
お問い合わせ：日本出版貿易株式会社 国内営業 1 部
〒101-0064 東京都千代田区猿楽町1-2-1 Tel: 03-3292-3767 Fax: 03-3292-8766
個人/学生購読について
NPG ネイチャー アジア・パシフィック　フルフィルメント & カスタマーサービス
〒162-0843 東京都新宿区市谷田町2-37 千代田ビル Tel: 03-3267-8759 Fax: 03-3267-8752
CHINA: www.natureasia.com/ch/subscribe; **KOREA:** www.natureasia.com/korea/subscribe

NATURE ONLINE

WWW.NATURE.COM/NATURE

CHIEF TECHNOLOGY OFFICER: Howard Ratner **ASSOCIATE DIRECTOR, NATURE.COM:** Daniel Pollock **PLATFORM DEVELOPMENT:** Jeremy Macdonald **WEB PRODUCTION:** Alexander Thurrell, Deborah Anthony **WEB DESIGN:** Juan-Carlos Sobrino, Luke Stavenhagen **PLATFORM INTERFACES:** Amanda Ward

SITE LICENCES, FULFILMENT & CUSTOMER SERVICES

SITE LICENCES: npg.nature.com/libraries feedback@nature.com
CHIEF OPERATIONS OFFICER: John Carroll **ASSOCIATE DIRECTOR OF OPERATIONS:** Dominic Pettit
CUSTOMER SERVICE: Liz Davidson **ORDER PROCESSING OPERATIONS:** Steve Poulter

NATURE PUBLISHING GROUP OFFICES

LONDON: The Macmillan Building, 4 Crinan Street, London N1 9XW Tel: +44 (0)20 7833 4000 Fax: +44 (0)20 7843 4596/7
BASINGSTOKE: Nature Publishing Group, Subscriptions, Brunel Road, Basingstoke, Hants RG21 6XS, UK Tel: +44 (0)1256 329242
WASHINGTON DC: 968 National Press Building, 529 14th St NW, Washington DC 20045-1938 Tel: +1 202 737 2355
NEW YORK: 75 Varick St, 9th Floor, New York, NY 10013-1917 Tel: +1 212 726 9200
BOSTON: 25 First Street, Suite 104, Cambridge, MA 02141 Tel: +1 617 475 9275
SAN FRANCISCO: 225 Bush Street, Suite 1453, San Francisco, CA 94104 Tel: +1 415 403 9027
MUNICH: Nymphenburger Strasse 14, D-80335 München Tel: +49 89 549057-13
PARIS: Moulin de Malicorne, Sainte Maure de Touraine, Paris 37800 Tel: +33 2 47 35 72 15
TOKYO: Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843 Tel: +81 3 3267 8751
INDIA: Nature Publishing Group, 3A, 4th Floor, DLF Corporate Park, Gurgaon 122002 Tel: +91 124 2881053/54

Nature® (ISSN 0028-0836) is published by Nature Publishing Group, a division of Macmillan Publishers Ltd (The Macmillan Building, 4 Crinan Street, London N19XW). Registered as a newspaper at the British Post Office. North and South American orders to subscriptions@natureny.com (Tel: +1 866 363 7860). Other orders to subscriptions@nature.com (Tel. +44 (0)1256 329242; Fax +44 (0)1256 812358). Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by *Nature* to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the relevant copyright fee is paid direct to CCC, 222 Rosewood Drive, Danvers MA 01923, USA. Identification code for *Nature*: 0028-0836/03. CPC PUB AGREEMENT #40032744. In the US, *Nature* (ISSN 0028-0836) is published weekly on Thursday, except the last week in December by Nature Publishing Group, 75 Varick St, 9th Fl, New York NY 10013-1917, USA. US Periodicals postage paid at New York NY, and additional mailing post offices. US POSTMASTER: send address changes to *Nature*, Subscription Dept, 75 Varick St, 9th Fl, New York NY 10013-1917, USA. Published in Japan by NPG Nature Asia-Pacific, Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843, Japan. © 2012 Macmillan Publishers Limited. All rights reserved.

I seek to move mountains.

MiSeq® Next-generation sequencing for all you seek.

You want amazing accuracy and performance on your benchtop. Illumina delivers—yet again. The MiSeq personal sequencer is the only fully integrated, truly end-to-end benchtop solution around. It's just one more example of why Illumina solutions generate a remarkable 90% of all the world's sequencing data. Discover what's possible.

www.illumina.com/miseq

illumina®

NEW ENGLAND
BioLabs® Inc.
enabling technologies in the life sciences
Stick together.

New DNA Ligases and Ligase Master Mixes

New England Biolabs offers the most extensive selection of high-quality and performance-optimized DNA ligases and ligase master mixes to streamline your cloning experiments.

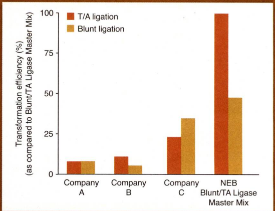
Our expanded portfolio now includes:

- Blunt/TA Ligase Master Mix, optimized for blunt-end and single-base overhang substrates
- Instant Sticky-end Ligase Master Mix, uniquely formulated for the rapid ligation of sticky-end substrates
- T7 DNA Ligase, specific for sticky ends
- ElectroLigase,™ directly compatible with electroporation

Stick together with DNA Ligases and Ligase Master Mixes from NEB.

To request a FREE SAMPLE of our new DNA Ligase Master Mixes, visit **NEBStickTogether.com**

Blunt/TA Ligase Master Mix outperforms the competition



Duplicate ligation reactions of blunt or T/A vector/insert pairs were set up according to the master mix vendors' suggestions. Equal amounts of ligated DNA were used to transform NEB 10-beta Competent E. coli (NEB #C3019) and triplicate plating was performed. Transformation results were averaged and graphed as a percentage of the highest performing reaction, T/A ligation using the Blunt/TA Ligase Master Mix.

ELECTROLIGASE™ is a trademark of New England Biolabs.

THIS WEEK

EDITORIALS

ERITREA The lessons of a decaying national science system **p.8**

WORLD VIEW Arab women know their rights and when to go home **p.9**

BIOSENSOR Blue is the colour when protein is present **p.10**

Murky manoeuvres

Scientific reform promised to give Italy's scientists the respect and autonomy they deserve, and political posturing must not be allowed to tip the burgeoning system off balance.

Three separate Italian court decisions, each in some way hinged on science, have shocked the international research community in recent months.

On 12 October, Italy's highest civil court ruled that compensation should be paid to a man who developed a tumour close to his brain that he claimed was caused by work-related use of mobile phones. On 22 October, a judge in L'Aquila sentenced six scientists and a government official to prison for manslaughter, saying that they failed to appropriately convey the risk of the 2009 earthquake, causing the deaths of 29 people who would otherwise have left their homes (see page 15).

The third decision, by a court in Brescia in July, ordered the temporary closure of Green Hill, a dog-breeding company in Montichiari that supplied animals for the toxicity tests officially required by bodies such as the European Medicines Agency and the US Food and Drug Administration, while mistreatment charges by animal-rights groups were investigated. The business had been regularly and rigorously checked by authorities over previous years, but has now effectively been destroyed because the judge placed the dogs in the care of the animal-rights groups, which distributed them to private homes.

Judges in Italy, as in democracies elsewhere, are supposed to make independent decisions based solely on the law. But the influence of a general societal mood is hard to avoid — and in Italy that society lacks understanding of, or respect for, science and its complexities.

Science is subject to a level of irrational suspicion in many countries, but in Italy there is a perception that science doesn't even matter — a state of affairs encouraged by decades of underfunding and political disdain. Italy invests just 1.26% of its gross domestic product in research and development (R&D), compared with Germany's 2.82% and a European Union (EU) average of 2%. In 2009, Italy employed only 226,000 full-time-equivalent R&D staff, whereas Germany had 535,000. The system has long suffered from the lack of a legally enforced meritocracy, allowing cronyism to taint academic appointment and promotion. Heads of research agencies have often been political appointees rather than competent experts.

Successive governments, well aware of the problems, introduced a series of reforms that tinkered with the system without fixing it, causing only further uncertainty. Then, three years ago, came a watershed: the reform-to-end-all-reforms intended to give more autonomy to research agencies, along with appropriate accountability. It sought to introduce an independent system to identify suitably qualified candidates as agency presidents (see http://doi.org/fwskwv), as well as a national research-evaluation agency whose assessments would be linked to funding. Designed by the centre-left government of Romano Prodi, it was finally passed into law in 2009 by the centre-right government of Silvio Berlusconi.

Enactment of such major reform has been a struggle, particularly for the newly appointed presidents of the 12 research agencies — which include the National Research Council, the National Institute of Nuclear Physics and the National Institute of Astrophysics — who are currently finalizing their new statutes. But a spirit of confidence has emerged. The agency presidents have formed a loose, cooperative alliance. And even the historically timid national academy, Accademia Nazionale dei Lincei, has become outspoken — for example, loudly challenging the L'Aquila court decision.

But research minister Francesco Profumo seems set on tipping things off balance again. In a murky manoeuvre, he announced reform plans in a financial newspaper on 11 October that would, along with other major changes, merge all 12 agencies into a single national organization — before the end of the year. He argued, unconvincingly and without a technical plan, that such a system would help to save money and win EU research grants. In the style of the old guard, whose day was thought to be done, he did not consult the general scientific community on the matter, not even agency presidents.

> ***"In Italy there is a perception that science doesn't even matter."***

It is impossible to imagine such a thing happening in, say, Germany, a country whose successful scientific system Profumo says he would like to emulate. German politicians and their administrations are in appropriate awe of their research-agency presidents and of the scientific culture they represent. It is also hard to imagine courts there crudely running rough-shod over science.

Profumo's amateurish proposal, which he tried to insert into Prime Minister Mario Monti's crisis-related financial law for 2013–15, did not survive first-round parliamentary scrutiny, but Profumo seems set to try to push for some sort of high-speed change — his government is slated to dissolve in March.

Crucial for now is that scientific leaders are left in peace to complete the reform-to-end-all-reforms, and that science doesn't fall victim, once again, to opaque politics. Building respect for science takes time. ■

Bad press

Japan's media have played a large part in exacerbating the effects of a fraud.

It is a shame that Shinya Yamanaka's recent Nobel prize had to be tainted by the shenanigans of Hisashi Moriguchi, the University of Tokyo project scientist who fabricated a story about having used Yamanaka's fêted technology on induced pluripotent stem (iPS) cells to treat patients who had heart failure.

The poor quality of journalism that led to the story being so widely reported was not an isolated incident in the coverage of science, in Japan or elsewhere. The *Yomiuri Shimbun*'s presentation of Moriguchi's 'accomplishments' was particularly disappointing. But other newspapers, including the *Nihon Keizai Shimbun*, have now admitted to having run unverified stories about Moriguchi over the past decade. Given the esoteric nature of the studies involved, reporting on science can be intimidating. So here are some practical steps to help a journalist challenge a specialist.

One can start by looking at publications. All scientists publish their results. If they don't it is a red flag. The publications give a scientist's affiliations, so if there is any doubt, it is easy to verify whether a scientist actually works where they say they do (a quick e-mail to Harvard University could have saved the *Yomiuri Shimbun* a lot of embarrassment). Publications also list the names of collaborators (making it easy to confirm with them that the scientist has done the experiments claimed), the names of the funders (making it easy to check whether resources were available) and declarations of conflicts of interest (revealing potential biases).

Most importantly, a journalist should talk to other researchers — those who do not collaborate with the scientist in question — about the study's significance and feasibility. Such researchers can usually be found by consulting references in the publication. If not — and an absence of proper references would be a warning sign — an Internet search will quickly bring up names. Although probably truer of North America and Europe than elsewhere, scientists are generally committed to keeping junk out of the literature. If it looks like junk, they'll tell you.

Of course, Moriguchi said that his latest results were not yet published. That should have triggered further questioning. Why would he present his results to the media first? Some scientists have a reason for doing so; Moriguchi did not. And it should have prompted a closer look at his work experience and past publications. Why do records of his career — which were available online — suggest that he had little or no experience in the field in which he was claiming to have made a revolutionary breakthrough? Why did he profess to work in a non-existent university Division of iPS Cell Research and Application? And why was he taking an unconventional and unfamiliar technology to the clinic? When he was questioned directly, as he was by *Nature*, things got worse. Why, for example, did he refuse to give the names of collaborators on the latest study? Poking the surface led to an outpouring of dubious statements.

People get away with fraud everywhere, but in Japan it seems that there are cultural factors that mean it goes unreported. Japanese scientists are less likely to be critical of their colleagues; there is less protection of whistleblowers who might not want to risk their careers; and journalists in Japan can be too polite, perhaps intimidated by the illustrious image carried by *sensei* and afraid to ask awkward questions. Possibly because of a lack of confidence in their English or because of differences between time zones, they often do not contact scientists overseas.

"Why would Moriguchi present his results to the media first?"

The situation is exacerbated by a recent Japanese epidemic: iPS-cell mania. With the excitement over Yamanaka's pioneering results, media outlets are rushing to get new iPS-cell stories first, sometimes regardless of their quality. This tendency is fuelled by a paranoid attachment to iPS-cell technology. Many news stories describe an international race to translate iPS-cell research into medical advances, which Japan might lose. This fear seems to have inspired Moriguchi, who lamented in 2009 that Japan was in danger of falling behind in iPS-cell research (in a Correspondence in *Nature*: H. Moriguchi and C. Sato *Nature* **457,** 257; 2009), and the *Yomiuri* reporters, who even envisaged a 'flexible' approval system in the United States that might allow Moriguchi to continue with his research.

This is all very silly. The beautiful thing about iPS-cell technology — and a major reason it won the Nobel prize — is that it can be used easily by scientists everywhere. If Japan wants to show its pride in Yamanaka's accomplishment, it should celebrate all achievements around the world. And if journalists want to understand just how important it is, they should put a new development in an international perspective. ■

Fight the power

Independence of academic institutions is crucial if nations are to rebuild.

It is said that truth is the first casualty of war. But after wars have ended and nations are rebuilding, truth is a potent foe, too.

A free press and a strong academic establishment pose great threats to a despotic regime, and are often the first to feel pressure. In a News Feature on page 24, we document such tensions in Eritrea, where there are efforts to remove foreign influences from the nation's medical schools and to break up its academic institutions. Eritreans who have fled or been exiled lament the sorry state of these institutions, and fear that by severing ties to US universities, the government has squandered chances of extraordinary gains in public health. (Representatives of the Eritrean government did not respond to several requests for comments on these and other allegations.)

Many who helped to lead Eritrea to its independence in 1991 and establish the current regime were academics, students and physicians — the very type of people (and, in some cases, the actual people) that the country's leaders are now marginalizing. This irony should not be a surprise, given that thriving academic institutions in new regimes can serve as hotbeds of dissidence, unrest and further revolution.

For a country such as Eritrea to reach a stable equilibrium, this destructive cycle must be broken. And this is perhaps the best lesson that Eritrea can give to nations that must build new governments. It is expected that academics will be involved in the overthrow of unjust regimes, and that they will subsequently be called on to support burgeoning governments. But these people must strive to remain independent from the start. The desire to trust and indulge former comrades may be strong, yet academics must reject all interference and resist any attempt by government to grab control of institutions.

To stand up to power in this way is difficult, but there is much at stake. Independent institutions will produce and support well-informed, independent citizens, who remain the most effective defence against corruption and the abuse of power. ■

ANNOUNCEMENT

Nature metrics

As of last week, *Nature* now provides a real-time online count of article-level metrics for its published research papers. Citation data, news mentions, blog posts and details of sharing through social networks, such as Facebook and Twitter, will be available for every research article published since last year.

Nature Publishing Group (NPG) hopes that the information will be of interest to readers, as well as feeding into the evolving debate about alternative ways to evaluate scientific output. The 2014 Research Excellence Framework exercise to assess UK research quality, for instance, will look at article citations and consider other measures for tracking research impact.

The information is available for 20 NPG journals published on nature.com, including the Nature research journals, *Nature Communications* and *Scientific Reports*. ■

MERCK MILLIPORE

Shaping Epigenetics Discovery.

Chromatin, DNA methylation, RNA analysis.

With EMD Millipore, epigenetics technology has never been more accessible or easier to use. From highly specific antibodies, to rapid bisulfite kits, to industry leading ChIP and next generation sequencing kits, we simplify epigenetics applications to allow you to design the most ambitious experiments to interrogate gene regulation.

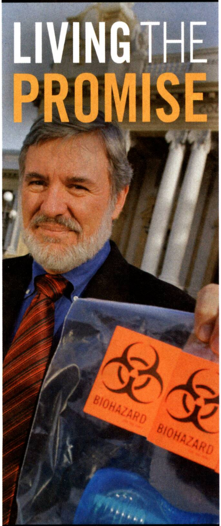
With our legacy of expertise from Chemicon and Upstate, we are committed to molding sophisticated yet simple solutions for understanding epigenetic regulation.

To learn more visit **www.millipore.com/EPIC**

Merck Millipore is a division of MERCK

Merck Millipore and the M logo are trademarks of Merck KGaA, Darmstadt, Germany. © 2012 EMD Millipore Corporation, Billerica, MA USA. All rights reserved.

UNIVERSITY OF CALIFORNIA
UCRIVERSIDE



Real World Solutions

Legally poisoned:
UCR Prof. Carl Cranor uncovers the alarmingly high rates of toxic chemicals that we are exposed to in everyday products and pushes for better protections in public health law.

Explore more impacts >
promise.ucr.edu

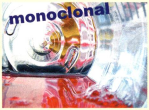
The Klein Bottle Opener

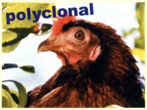
A geometrical joke, 3D printed in stainless steel. The perfect touch for any math fan's kitchen! For this plus more math & science art, come to www.bathsheba.com

Yes, it really opens bottles.
Now available inscribed!

Davids Biotechnologie
Antibodies made in Germany
Your manufacturer since 1996

Davids GmbH Biotechnologie





www.davids-bio.de

nature

A key thread in your scientific career

Whether in print, online or mobile, *Nature* is always at the forefront of science and is your forum to read, watch, listen and engage with key research, news and opinion.

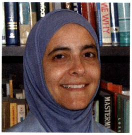
How will you access the next landmark discovery?

NATURE.COM/NATURE

nature publishing group npg

WORLD VIEW

A personal take on events



A. AWAD

How women scientists fare in the Arab world

Rana Dajani *argues that true equality for women scientists requires recognition of their family roles as well.*

I will never forget a rich Arab businessman I met in New York at the Clinton Global Initiative. He had an American lady in charge of his non-governmental organization. She was a great woman dedicated to her work, who would stay until midnight doing her job. He said that women in the Arab world will never succeed until they are ready to stay at work until midnight like the American woman. He complained that Arab women want to go home at five o'clock to look after their families. I was upset. I told him that it should be their choice how long they want to work; and if they have priorities such as family and want to leave at five, then this should be respected. It is men like him, and in some cases women, who undermine who we are.

The issues faced by women in science are beginning to receive more attention. On this page last week, Athene Donald highlighted one initiative to tackle gender bias (A. Donald *Nature* **490,** 447; 2012). I would like to offer a perspective from the Islamic world.

Despite the impression given by extremists, Islam gives women the right to education. More than four in ten women who go to university in Jordan go into science, engineering or medicine. Women outnumber men on courses in natural science, pharmacology and agriculture; numbers are equal in maths and computer science; and one in three engineering students in Jordan is a woman.

Some of the problems faced by women scientists in the Middle East are the same as those faced by women around the world. Our productivity, for instance, is measured on a male scale. The years we spend taking care of children are not calculated as part of the gross domestic product of a country. What is more important — to build physical things or to nurture a human being?

ONE MUST NOT **FALL INTO THE TRAP** OF TRANSFERRING SOLUTIONS FROM **ONE CULTURE** TO ANOTHER.

As an example of this male scale, L'Oréal and the United Nations Educational, Scientific and Cultural Organization are running a competition to award fellowships to Arab women scientists — but you have to be under 40 to enter. This is biased, and based on metrics from a male-dominated world, in which if a man doesn't make it by 40 he is a failure.

The feminist movement was a good thing, but it was too focused on equality with men and failed to enable us to respect ourselves as women and to be proud of who we are.

Another common challenge to all women scientists is lack of mentoring and networking. Most women scientists everywhere have two jobs — work and home — and most will not give up home for work. They will always be worried about the children, want to be with them, and feel that the father's presence won't compensate for their own absence. So they don't take time after work to have a coffee with their colleagues.

Yet this informal environment is where scientists learn what is going on; where they lobby, network, mentor and get mentored. Women don't have the time. Networking is an extra effort. Men mentor each other and spend time together after work, fostering the men's club. Women rush home to take care of children, not because they have to but because they want to.

➲ **NATURE.COM**
Discuss this article online at:
go.nature.com/dbklie

This is a major obstacle for women scientists in terms of opportunities, learning and support. That is why mentoring projects — something we lack in the Arab world — are important. But social media allow mentoring online, and some women scientists now plan to start an online mentoring scheme for women scientists in Jordan, in collaboration with the country's first woman university president, Rowaida Maaitah.

Women also have challenges specific to the Middle East. These may not be so obvious because they are subtle, and must be identified, studied and solved by Arab women themselves. For instance, the September study about a bias among US scientists against women, mentioned by Donald last week (C. A. Moss-Racusin *et al. Proc. Natl Acad. Sci. USA* http://doi.org/jkm; 2012), would not necessarily translate to the Arab Muslim world, where the prevailing attitude among both men and women is that women work hard and are more dependable than men. One must not fall into the trap of transferring solutions from one culture to another.

I know of an American researcher who went to Bulgaria to help women fight for their rights. She went assuming that they would want to demand to work. But Bulgarian women who had lived under Communism wanted the exact opposite. They wanted the freedom to stay at home if they chose.

A much-misunderstood issue is the covering of the hair and sometimes the face by Muslim women. In the West, this is often considered a sign of oppression. Yet more than half the female students and academics in the Arab world choose to cover their hair for religious reasons, compared with fewer than 10% 20 years ago. These young women are educated, affluent and independent. I have a graduate student who covers her face who told me that she believed she would win a Nobel prize. This is not oppression.

I see the Arab spring as an opportunity for women to learn about their rights and to advocate for them — to distinguish between what is tradition and what is religion. This would weed out extremists who, through ignorance, distort the image of Islam. And it would weed out opportunists who want to misrepresent Muslim women. Throughout the Islamic civilization that flourished in the Middle Ages there were more than 8,000 women scholars. There are many more on the way today. ■

Rana Dajani *is assistant professor of molecular biology at the Hashemite University in Zarqa, Jordan, and Fulbright visiting professor at Yale University.*
e-mail: rdajani@hu.edu.jo

RESEARCH HIGHLIGHTS

Selections from the scientific literature

CANCER

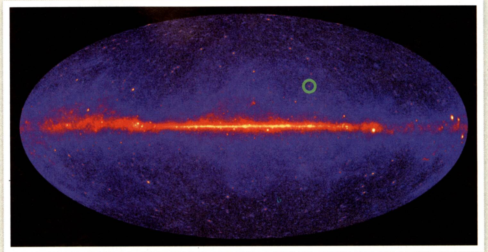
Predictor pairs images and genes

Automated image processing can be integrated with molecular profiling to provide a fuller portrait of cancer.

Pathologists routinely use visual examinations of cell types in tumour biopsies to direct patient care. However, it is hard to integrate these visual analyses with data from gene-expression studies.

Florian Markowetz and Yinyin Yuan at the University of Cambridge, UK, and their team came up with software that can analyse images of stained tissue sections to determine the identity and arrangement of cells in tumours. For some cell types, certain spatial patterns were associated with longer patient survival. However, an algorithm that combined image-based and gene-expression data predicted survival more accurately than algorithms that used either type of information alone.

Sci. Transl. Med. 4, **157ra143 (2012)**



AEI/FERMI LAT COLLABORATION/DOE/NASA

ASTRONOMY

Picking out a predatory pulsar

Researchers have used raw computing power to hunt down a 'black widow' pulsar that is evaporating its companion star.

Pulsars are stellar remnants that emit lighthouse-like beams of radiation. They often emit gamma rays, but can usually be spotted only if they also emit easier-to-detect radio waves. However, Holger Pletsch of the Max Planck Institute for Gravitational Physics in Hannover, Germany, and his colleagues found the current pulsar (pictured; circled) through a blind search of data from the Fermi Gamma-ray Space Telescope.

Using computers to analyse huge swathes of raw data, the team picked out the pulsar, which takes 93 minutes to orbit its companion star. This orbital period is the shortest of any binary pulsar of this type yet found.

Science **http://dx.doi.org/10.1126/science.1229054 (2012)**

MICROBIOLOGY

Bacteria beaten by bacteria

Infections caused by a strain of *Clostridium difficile* responsible for recent epidemics might be treatable using a mix of gut bacteria.

A team led by Trevor Lawley of the Wellcome Trust Sanger Institute near Cambridge, UK, infected mice with the *C. difficile* strain and then treated them with an antibiotic commonly used in humans. Instead of killing the pathogen, the antibiotic displaced other gut bacteria and permitted a persistent *C. difficile* infection.

Treating the infected mice with faeces from healthy animals restored their intestinal flora and resolved their infections. A mixture of six different bacterial species isolated from the faeces had the same effect, promising a treatment more palatable to patients than faecal transplantation.

PLoS Path. 8, **e1002995 (2012)**

CLIMATE-CHANGE ECOLOGY

Plankton diversity loss looms

They are responsible for about half of all photosynthesis on Earth — and plankton could be drastically affected by climate change.

Mridul Thomas and his team at Michigan State University in East Lansing considered 194 phytoplankton strains. Using the existing literature, the authors estimated the maximum growth rate, optimum temperature for growth and the temperature range over which growth can occur, for each of the strains. Many strains seem to be tightly adapted to the average temperature at their location. Tropical strains, in particular, tend to have optimal growth temperatures at or just below the mean temperature in their environment. The authors' models indicate that an average temperature rise of just 2 °C in the tropics by 2100 could reduce the diversity of phytoplankton in the region by a third — unless, that is, the plankton can evolve greater heat tolerance.

Science **http://dx.doi.org/10.1126/science.1224836 (2012)**

BIOSENSORS

Naked-eye ELISA developed

Researchers have developed a sensitive bioassay that can be read with the naked eye.

The assay is based on a laboratory technique known as an ELISA (enzyme-linked immunosorbent assay), in which an enzyme generates a coloured compound whenever

R. DE LA RICA QUESADA/STEVENS GROUP

antibodies recognize a target molecule. However, detecting low levels of this compound requires expensive instruments that few labs in the developing world can afford, so Molly Stevens and Roberto de la Rica of Imperial College London devised an alternative.

In their method, the ELISA enzyme controls the aggregation of nanoparticles, giving rise to a blue colour if a target protein is present and a red colour if it is not. Although the bioassay cannot quantify protein levels, it can detect an HIV protein at concentrations as low as 1 attogram per millilitre.

Nature Nanotechnol. **http://dx.doi.org.10.1038/nnano.2012.186 (2012)**

MICROBIOLOGY

Cheating yeast finish last

The 'tragedy of the commons' holds that cheaters have an advantage over cooperators because cheaters benefit from common goods without contributing to them. Studies in yeast suggest a new mechanism to avert such a tragedy.

Adam James Waite and Wenying Shou of the Fred Hutchinson Cancer Research Center in Seattle, Washington, started cultures that had equal amounts of three yeast strains: one that produced adenine and required lysine; another that produced lysine and required adenine, and a third, 'cheating' strain, which required lysine but did not supply any nutrients.

Contrary to expectations, 'cooperative' strains dominated in some cultures, and could occasionally drive cheaters to extinction. Genome sequencing revealed that the dominating strains had adapted to their new environment by accumulating mutations that improved nutrient transport. When these mutations arose in cooperative strains and compensated for the cost of cooperation, the cheaters were outcompeted.

Proc. Natl Acad. Sci. USA **http://dx.doi.org/10.1073/pnas.1210190109 (2012)**

CLIMATE

Tailored geoengineering

Climate-engineering techniques that cool Earth by reflecting sunlight back into space may be tailored to minimize negative effects on individual regions without compromising overall cooling.

Douglas MacMartin at the California Institute of Technology in Pasadena and his colleagues used a global climate model to explore the impact of techniques such as the injection of aerosols into the stratosphere. The team then modelled the effects of varying the interventions spatially and seasonally, and showed that the average global temperature could be reduced while still supporting goals such as the recovery of Arctic sea ice.

The model suggests that climate interventions could provide the world with more than a single 'global thermostat', the authors say.

Nature Clim. Change **http://dx.doi.org/10.1038/nclimate1722 (2012)**

COMMUNITY CHOICE

The most viewed papers in science

ANIMAL BEHAVIOUR

Bears show knack for numbers

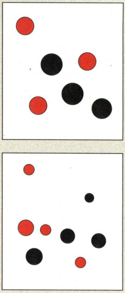
HIGHLY READ
on www.journals.elsevier.com/animal-behaviour over the past 3 months

Bears may be able to estimate and compare numbers of items.

Jennifer Vonk of Oakland University in Rochester, Michigan, and Michael Beran of Georgia State University in Atlanta trained three captive black bears (*Ursus americanus*) to distinguish between two groups of dots (pictured) displayed on a touch-screen computer. One bear was trained to touch the group containing the most dots, and the others to select the group with the fewest. By varying the size and positions of the dots, researchers tested whether the animals could recognize dot number independently of the total area covered by the dots. Although the bears showed a preference for dots occupying a larger area, they also showed some ability to judge the relative numbers of dots.

The bears' numerical skills may have evolved to help them in complex foraging environments, the authors suggest.

Anim. Behav. 84, **231–238 (2012)**



J. VONK/ANIM. BEHAV./ELSEVIER

CONDENSED MATTER PHYSICS

Building a space-time crystal

Just as a crystal consists of a regular array of particles repeated in space, so a space-time crystal should consist of a regular pattern of particles that also repeats cyclically over time. Now, Xiang Zhang at the University of California, Berkeley, and his team have published the first proposal for an experiment that could realize this abstract notion.

The authors' idea is to trap a ring of cold ions in a magnetic field so that they adopt a regular arrangement and then rotate in their lowest energy state, thereby creating temporal repetition. Such a device might be able to store quantum information, and could thus have applications in quantum computing, the researchers say.

Phys. Rev. Lett. 109, **163001 (2012)**

PALAEONTOLOGY

Palaeoflamingo nest found

A fossilized nest, found in Spain and containing five eggs, belonged to a previously unknown species of palaeoflamingo, the ancient ancestor of the modern, long-legged bird. The nest — made from twigs and leaves 15 million to 20 million years ago — was found alongside bone fragments, encased in limestone in the Bardenas Reales de Navarra Natural Park.

Gerald Grellet-Tinner from the Field Museum in Chicago, Illinois, and his team report that the eggshells are characteristic of flamingos, whereas the nest and the number of eggs more closely resemble those of grebes, freshwater diving birds. Modern grebes and flamingos differ in their nest-building and feeding styles, but DNA studies have suggested that the two species are closely related. The present discovery supports that connection and points to a time when the two species shared survival strategies.

PLoS ONE 7, **e46972 (2012)**

NATURE.COM
For the latest research published by *Nature* visit:
www.nature.com/latestresearch

SEVEN DAYS *The news in brief*

ERTIFIED ELEVATOR

KEYSTONEUSA-ZUMA/REX FEATURES

Titanic storm batters US east coast

Hurricane Sandy, which was downgraded to a post-tropical cyclone as wind speeds dropped from about 145 kilometres per hour to around 105 kilometres per hour, made landfall near Atlantic City, New Jersey, at about 8 p.m. Eastern time on 29 October. The storm knocked out power to millions and closed public transport in numerous cities, including New York, which was among the hardest hit. More than 15,000 flights have been cancelled, and estimates of the damage run to well over US$10 billion. See tinyurl.com/99wwz9w for more.

POLICY

Verdict shock wave

The six-year manslaughter sentences handed to six scientists and a government official for advice they gave preceding the 6 April 2009 earthquake in L'Aquila, Italy, have triggered concerns worldwide over the future of science advice. Anne Glover, chief scientific adviser to the European Commission, is among scientists warning that the verdict could mean researchers are less likely to agree to advise governments. See page 15 for more.

Electronics boost

India unveiled an electronics policy that aims to build a domestic chip-manufacturing industry, creating 2 million jobs by 2020. Proposed initiatives include a fund to promote electronics research and development, and an institute focusing on semiconductor chip design. The government also plans to bolster postgraduate education to produce about 2,500 PhDs annually in electronics by 2020.

Biofuels warning

Pushing up the proportion of alga-based biofuels used for transport to 5% of the total US requirement "would place unsustainable demands on energy, water, and nutrients", says a report from the US National Research Council. Consumption of water and fertilizer for the cultivation of algal biofuels can be prohibitive, and such concerns need to be addressed if these fuels are to fulfil their promise, adds the report, released on 24 October. See go.nature.com/kyptgk for more.

Nuclear restart

China will start approving new reactors again, ending a 19-month ban triggered by the nuclear disaster in Fukushima, Japan, the Chinese government announced on 24 October. But enhanced safety standards and a temporary ban on inland reactors, which account for one-third of those planned, mean that nuclear reactors will supply 40 gigawatts of power by 2015, rather than 50 gigawatts as planned. China currently has 15 reactors supplying 12.5 gigawatts, or 1.8%, of its power.

Fisheries trade-off

A proposal that would have placed the European Union's fisheries policies on a sounder scientific footing was watered down at a meeting of ministers on 23 October. Conservationists have criticized the decision to spend money from a new European Maritime and Fisheries Fund on modernizing fishing fleets, arguing that such measures will not improve the sustainability of fish stocks and that more funds should be spent on data collection and monitoring fishing fleets. See go.nature.com/t6j4qc for more.

BUSINESS

Re-enter the Dragon

A commercial resupply capsule has returned safely from the International Space Station. The Dragon spacecraft, built by SpaceX of Hawthorne, California, splashed down some 400 kilometres off the California coast on 28 October. It was carrying around 750 kilograms of return cargo, including scientific samples. This is the first of at least 12 commercial resupply missions that SpaceX will send to the station as part of a US$1.6-billion contract.

Solar exit

German engineering company Siemens announced that it would sell its solar business on 22 October. The Munich-based company also said that it was ending its participation in DESERTEC, a project seeking to produce power by tapping solar energy in the Sahara and other deserts. See page 16 for more.

Nuclear rights

Hitachi has bought the rights to build nuclear power plants in the United Kingdom. The Japanese firm announced the purchase of the Horizon nuclear project from German owners E.ON and RWE for a reported £700 million (US$1.1 billion). The deal clears the way for Hitachi to seek licensing approval for up to three 1,300-megawatt advanced boiling water reactors at each of two sites in Anglesey and Gloucestershire.

PEOPLE

Synchrotron chief

Three years after converting its 3-kilometre-long linear particle accelerator into an X-ray laser, the SLAC National Accelerator Laboratory in Menlo Park, California, named Chi-Chang Kao as its director, on 24 October. Kao, who develops applications of synchrotron radiation and was an associate laboratory director at SLAC, is the first scientist who is not a particle physicist to lead the facility. Having transformed itself from a particle-physics stronghold into an X-ray light source, SLAC now draws structural biologists, material scientists and other researchers.

RESEARCH

Starry, starry night

I. TOLEDO, M. KORNMESSER/VVV CONSORTIUM

The European Southern Observatory (ESO) unveiled a 9-gigapixel image of the centre of the Milky Way (**pictured**) on 24 October. If printed at the standard resolution of a book, the image, from the ESO's VISTA infrared survey telescope in Chile, would be 9 metres long and 7 metres high. Astronomers used the telescope data to create a catalogue of 84 million stars — the largest such catalogue compiled to date.

Radiation study

The US National Academy of Sciences is undertaking a pilot study to look for cancer risks around six nuclear power plants and a nuclear fuel facility. The study was called for by the Nuclear Regulatory Commission in response to long-standing public concerns about radiation affecting the health of people living around the plants (see *Nature* **472,** 15; 2011). Announced on 23 October, the academy study is expected to continue into at least 2014 and will cost around US$2 million.

Coastline review

An eight-year survey of coastal regions by China's State Oceanic Administration was completed on 26 October. Of 52 cities surveyed, 28 have serious water shortages, according to the assessment, which included 19,057 kilometres of continental coastline and 10,312 islands.

Shanghai telescope

China unveiled a radio telescope at the base of Sheshan Mountain in Shanghai on 28 October. The telescope will be used with three others across China for very-long baseline interferometry, a technique that combines data from different telescopes to produce images of higher resolution than any of the telescopes can provide alone.

COMING UP

2–3 NOVEMBER
A Cancer Research UK Cambridge Research Institute symposium will focus on unanswered questions in cancer sequencing.
tinyurl.com/9ns9c5n

4–7 NOVEMBER
Updates from the Mars Curiosity rover will be discussed at the Geological Society of America meeting in Charlotte, North Carolina.
tinyurl.com/9ouo7kr

5–7 NOVEMBER
The Fourth Canadian Science Policy Conference will discuss issues including how fundamental research can drive innovation.
www.cspc2012.ca

7–8 NOVEMBER
The European Food Safety Authority gathers scientists in Parma, Italy, to discuss challenges in risk assessment.
tinyurl.com/9zy5vgt

EVENTS

Canada quake

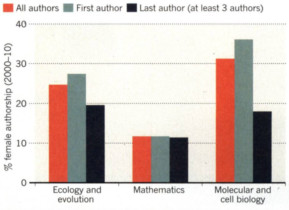
An earthquake off the west coast of Canada on 27 October triggered tsunami warnings for the Pacific. The magnitude-7.7 quake, centred close to the Queen Charlotte Islands, caused no major damage, however, and ultimately produced waves of only about a metre in Hilo, Hawaii. See go.nature.com/2ug81s for more.

TREND WATCH

An analysis of the gender of authors in research articles shows that women are not only under-represented as authors, but are also much less likely to be in the prestigious position of last author in all surveyed science fields except mathematics (see chart), in which authors tend to be listed alphabetically. The study, based on millions of articles in the academic journal archive JSTOR, was led by researchers at the University of Washington in Seattle and New York University (see www.eigenfactor.org/gender).

GENDER BIAS IN RESEARCH

Women are under-represented as authors on research papers — but especially as last authors.



SOURCE: J. WEST/UNIV. WASHINGTON

NATURE.COM
For daily news updates see:
www.nature.com/news

GOT AN IDEA BREWING?
MAKE IT HAPPEN WITH A CATALYST GRANT OF UP TO $25k

If you've got an idea that blends science with technology, we've got the funding and resources to make it happen.

Get in touch at **digital-science.com/catalyst**

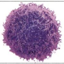
DIGITAL science

NEWS IN FOCUS

ANIMAL RESEARCH An urgent housing search for the NIH's retired chimps **p.18**

PHYSICS Quantum theorists make light work of space-time rips **p.19**

IMMUNOLOGY Chinese vaccine tackles neglected form of hepatitis **p.21**



GENOMICS Single-cell sequencing reveals hidden diversity **p.27**

F. MONTEFORTE/AFP/GETTY

Official reassurances convinced many residents of L'Aquila that the earthquake risk was low, a factor that many blame for the 2009 quake's high death toll.

SCIENTIFIC ADVICE

L'Aquila verdict row grows

Global backlash greets sentencing of Italian scientists who assessed earthquake risk.

BY NICOLA NOSENGO IN ROME

To scientists in Italy, offering advice on risks looks to be a hazard in itself. In the aftermath of a judge's decision to sentence a group of seismic-risk experts to six years in prison, researchers in the country are demanding that legal safeguards be put in place for advisers, and that a clear division of responsibilities be made between scientific experts and government decision-makers.

On 22 October, a court in L'Aquila found six scientists and one government official guilty of playing down earthquake risks in the region just days before a devastating earthquake hit the city on 6 April 2009. Prosecutors had argued that the group's assessment of a surge in seismic activity, and a subsequent press conference involving two members of the group, focused on simply reassuring the public rather than providing a careful evaluation of the potential hazards. Residents of L'Aquila say that they typically leave their homes when swarms of tremors shake the region, but the calming messages from the experts, the court heard, persuaded dozens of people to stay indoors. Many then perished as numerous buildings collapsed during the magnitude-6.3 quake (see *Nature* **477**, 264–269; 2011).

➲ NATURE.COM
Read more about the background to this case:
go.nature.com/mkvy3o

The severity of the sentences, however, surprised many — even the public prosecutor, Fabio Picuti, who had sought terms of four years — and condemnation quickly followed from the global scientific community (see *Nature* **490,** 446; 2012).

Stefano Gresta, president of Italy's National Institute of Geophysics and Volcanology in Rome (where two of the convicted scientists worked) said at the time that the verdict "is likely to compromise the right/duty of scientists to take part in the public debate" because they will be "in fear of being convicted". Statements of support for the scientists came from the American Geophysical Union, the Geological Society of America, the International Human Rights Network of Academies and ▶

▶ Scholarly Societies and the science-promotion group Euroscience. The US National Academy of Sciences and the Royal Society, UK, issued a joint statement saying that the verdict "could lead to a situation in which scientists will be afraid to give expert opinion for fear of prosecution or reprisal". And the chief scientific adviser to the European Commission, Anne Glover, told *Nature* that she is concerned about the fallout from the case. "It may become more difficult to motivate scientists to advise on inherently uncertain issues," she says.

The verdict will set no legal precedent abroad. But there is a risk that L'Aquila could become an "informal precedent, in that it breaks down an aura of immunity surrounding science", says Sheila Jasanoff, who specializes on the intersection of science and the law at Harvard Kennedy School in Cambridge, Massachusetts. "People and lawyers may be more willing to sue when they feel bad advice has been given."

The day after the verdict, Luciano Maiani resigned as President of Italy's major risks commission, the expert panel that advises the national government on environmental hazards. Most of the convicted scientists had given their assessment of the seismic risk at L'Aquila as serving members of the committee. Maiani says that he was protesting not the verdict itself but the lack of legal protection for his committee.

Maiani, a physicist at the University of Rome, La Sapienza, and former director-general of CERN, Europe's particle-physics facility near Geneva in Switzerland, says that, since taking up his post at the beginning of this year, he has been asking the government to commit legal assistance to the committee members, as well as financial insurance in case of civil proceedings, but his requests have been refused. "This way it is impossible to give truly independent advice," he argues. "The obvious consequence is that scientists will tend to be more conservative in their advice, like some medical doctors who order too many tests and surgeries for fear of being sued. But this is not in society's best interest."

The full reasoning behind the verdict is not yet public, leaving some Italian legal experts baffled as to why the judge issued such a heavy sentence, making no distinction between scientists in an advisory role and government officials entitled to take decisions and speak to the public. "Normally, the official who asks for expert advice remains legally responsible for the actions he then takes, unless the advice was wilfully and seriously wrong," says Stefano Rodotà, a legal specialist at La Sapienza and a former member of the Italian Parliament.

> ***"Scientists are not elected, so it cannot be up to them to decide how to deal with an emergency."***

Glover says that the case underlines the importance of having clear rules on how scientific advice is given and used. "Before a scientist gives advice, it should be crystal clear to him or her what someone else is going to do with it", she says, but, ultimately, elected officials should be held responsible for decisions resulting from this process. "Scientists are not elected, so it cannot be up to them to decide how to deal with an emergency," Glover says.

Many scientists contacted by *Nature* agree that better legal protection, along with transparent guidelines about the obligations of science advisers, are long overdue in Italy. "The case resulted from the fact that the legal role of scientific advisers is still not well defined in Italy," says Mariachiara Tallacchini, who studies science-related legal issues at the Catholic University of Piacenza. "Countries such as the United Kingdom and the United States are more advanced in regulating science policy".

John Beddington, the UK chief scientific adviser, agrees. "I do not think such an outcome would be possible in the United Kingdom, unless the advice was demonstrably grossly negligent or wilfully malicious," he told *Nature*. "And in the case of civil proceedings, all advisers are indemnified by government." Similar protection is granted to science advisers in the United States, where seismologists advising national and state governments would be immune from such prosecution.

In Italy, there may be no time for reform before the next crisis. On 26 October, a magnitude-5.0 earthquake hit the Pollino region in the south of the country, where shocks have been going on for months — a situation very similar to that seen in L'Aquila in 2009. Maiani, who has not yet left the major risks commission, says that he and the other committee members will continue to serve during the current emergency. ■ SEE EDITORIAL P.7

ENERGY

Sahara solar plan loses its shine

Siemens' decision to pull out of DESERTEC reignites doubts.

BY DEVIN POWELL

Dimming prospects for solar energy have caught up with a massive renewable-energy project planned for the Sahara Desert. By 2050, according to its backers, DESERTEC, a network of solar plants and other renewable sources scattered across North Africa and the Middle East, could generate more than 125 gigawatts of power that could be used locally or delivered to Europe through high-voltage direct-current cables beneath the Mediterranean Sea. But one of its major backers, Siemens, based in Munich, Germany, now says that it will leave Dii, the consortium trying to advance DESERTEC, by the end of the year.

"We see our part in Dii as done," says spokesman Torsten Wolf of Siemens, one of 13 founding partners of the consortium, which is also based in Munich.

Siemens also said that it will pull out of the solar-energy business altogether. Its decision was made in response to falling government subsidies for solar energy and a collapse in the price of solar equipment. But to DESERTEC'S critics, Siemens' exit also adds to doubts about the plan, which is expected to cost hundreds of billions of dollars. "DESERTEC is an ambitious attempt to do everything at once," says Jenny Chase, an analyst at Bloomberg New Energy Finance in Zurich, Switzerland. "I think it's something that will be achieved organically, bit by bit, which will probably be cheaper, easier and achieve the same results."

DESERTEC's origins lie with retired particle physicist Gerhard Knies, who, after the 1986 Chernobyl nuclear disaster, had the idea of harvesting the Sahara's plentiful sunshine for energy. With the help of Prince El Hassan Bin Talal of Jordan, Knies brought together research institutes in Germany and North Africa, including some in Morocco, Algeria and Egypt, to start looking into the idea.

"For generating electricity from renewable, carbon-free sources, the economics makes more sense if we go to the Middle East and North Africa," says Ernst Rauch, Dii representative for the insurance company Munich Re, one of the consortium's shareholders.

Siemens provided funds and technical expertise for preliminary studies, and in June, Dii published the result: a report that maps out the most cost-effective distribution of renewable-energy sources in 2050, based on simulations run by the Fraunhofer Institute for Systems and Innovation Research in Karlsruhe, Germany (see 'Power play').

Paul van Son, Dii's chief executive, says that

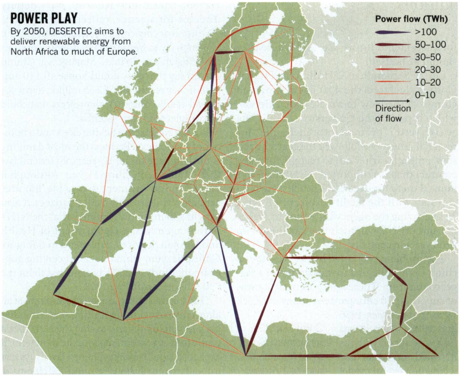
Solar thermal power plants offer one option for supplying Europe with electricity through the DESERTEC project (artist's concept).

DESERTEC/BRIGHTSOURCE

he is not concerned about losing Siemens for the next phase of the work, a detailed look at specific projects. "It will not really affect us," he says, pointing out that Siemens is only one of dozens of shareholders and partners.

Siemens' exit from solar energy is an about-face for the company, which had been investing in solar thermal energy, long considered a key feature of DESERTEC, after acquiring the solar-thermal equipment designer Solel, based in Beit-Shemesh, Israel, in 2009. Plants built with this technology use mirrors to concentrate sunlight on material that can absorb its heat. When released, the heat boils water, creating steam that can drive an electricity-generating turbine.

Solar thermal plants have become increasingly hard to sell in recent years, says Wolf, owing to the falling price of a competing technology, silicon solar panels. Between 2006 and 2012, the cost of these photovoltaic panels fell by around 65%, partly as a result of a glut of solar cells on the market. Waning government subsidies for solar installations have made staying in the business even harder by lowering the demand for solar power, says Matthew Feinstein, an analyst at Lux Research in New York.



POWER PLAY

By 2050, DESERTEC aims to deliver renewable energy from North Africa to much of Europe.

"It's been a bloodbath for the past year or so," says Feinstein. Last December saw the bankruptcy of Berlin-based Solon, Germany's first publicly traded solar company. Q.Cells, headquartered in Bitterfeld-Wolfen, one of the world's top solar-cell producers, followed in March. Siemens' exit from the business will deepen the gloom.

But Thiemo Gropp, director of the nonprofit DESERTEC Foundation that set up Dii, says the fall in costs that drove Siemens out of the business will ultimately benefit the project. "Other companies will fill this gap with their own products," he says. DESERTEC has endorsed one such project, a solar thermal plant planned for a site in Tunisia. Nur Energie, the London-based company behind the undertaking, hopes to supply Italy with power from the plant through an underwater cable.

And Siemens' relationship with DESERTEC may not be entirely over. The company says that it still supports the mission in principle, and may have something to offer down the road. Wind energy, for example, has taken on an increasingly important role in plans for DESERTEC, which hopes to tap into Africa's coastal winds. And Siemens now plans to focus on its wind portfolio.

The company has already received its first turbine orders from Africa. Morocco, which hopes to generate 6 gigawatts of power annually from renewable sources by 2020, bought 44 or them, to be installed at two wind farms. Although the projects have not been officially endorsed by DESERTEC, Wolf says that they fit the spirit of the endeavour perfectly. ■

Animals at Chimp Haven in Louisiana, where activists would like retired NIH research chimps to be sent.

A. FULTZ/CHIMP HAVEN

ANIMAL RESEARCH

NIH faces chimp housing quandary

Dozens of chimpanzees retired from research may have to continue to live in lab-like conditions.

BY MEREDITH WADMAN

It is not easy to find living space for a great ape at short notice, let alone more than 100 of them. Yet that is precisely the problem that administrators at the US National Institutes of Health (NIH) are scrambling to solve, as the biomedical agency takes its most visible and decisive step away from invasive research on chimpanzees.

Scrutiny of the NIH's chimp research enterprise has been intensifying since the release last December of an Institute of Medicine report, which declared most of the invasive chimp studies to be scientifically unnecessary (see *Nature* **480,** 424–425; 2011). The agency, based in Bethesda, Maryland, immediately put a moratorium on new grant applications for work involving chimps. In January 2013, a working group will recommend which of the grants already in progress should continue to be funded. The group will also advise on how many research-eligible chimps the agency should maintain for current and future use, and where they should be housed.

On 21 September, NIH director Francis Collins declared the 110 agency-owned chimps at the New Iberia Research Center, which is part of the University of Louisiana at Lafayette, "permanently ineligible" for research. The move followed the centre's decision a month earlier not to reapply for a key contract that has supported the NIH chimps housed there for decades. The existing NIH contract with the facility expires in August 2013, leaving the agency little time to avert a housing crisis for animals that can live for up to 60 years in captivity.

The problem presented by the New Iberia chimps is just the first manifestation of a bigger conundrum. The federal government owns or supports 670 chimpanzees, many of which were bred between 1986 and 1995, when it was hoped — incorrectly, as it turned out — that they would be a useful model for HIV/AIDS. Although some have been used in virology studies and in the development of monoclonal antibodies, their use by federal researchers now looks set to dwindle.

➲ NATURE.COM
For more on US chimp research, see:
http://doi.org/fgkmc9

Critics say that the housing problem should have been addressed long ago. "This is emblematic of the NIH's failure to plan," says Eric Kleiman, a research consultant for the Animal Welfare Institute in Washington DC. "The writing has been on the wall for how many years now?"

Collins announced the withdrawal of the 110 chimps from research in personal calls to British primatologist and chimp-welfare activist Jane Goodall, and Wayne Pacelle, director of the Humane Society of the United States in Washington DC. They were pleased with the news, but not with the NIH's plans for housing many of the chimps. Collins said he would move 10–20 animals to fill available space at Chimp Haven in Keithville, Louisiana, the only federally supported chimpanzee sanctuary. The rest would go to the Texas Biomedical Research Institute in San Antonio, where each social group of four to six chimps would be housed in an indoor–outdoor enclosure about the size of a squash court, with extra space for elevated perches.

Chimp advocates say that Chimp Haven, a forested 80-hectare refuge that is currently home to 124 chimps, could accommodate the animals in more appropriate conditions than the research institute. "There is no comparison between a place like Chimp Haven and Texas Biomed," says Kleiman. "Chimp Haven is chimpanzee-centred. Texas Biomed is a lab. It's caging."

"In a perfect world, we would absolutely like to move all of the chimps directly to Chimp Haven," says Kathy Hudson, NIH deputy director for science, outreach and policy. "We are working collaboratively with Chimp Haven to try to figure out what are the options for being able to do that." Managers at Chimp Haven say that they could house all 110 animals if they received US$2.55 million to pay for shovel-ready construction projects that could be completed in four months.

But the NIH faces a ticking clock and a number of roadblocks. Perhaps the most daunting is the language of the 12-year-old federal law that established Chimp Haven. Although it obliges the government to provide 'lifetime' care for retired research chimpanzees, it also caps at $30 million the money that the NIH's parent agency, the Department of Health and Human Services, can spend in doing so. Chimp Haven, which began receiving government funds in 2002, is expected to hit the $30 million cap during 2013.

Hudson says that the agency is looking at all alternatives, including finding space in other private sanctuaries and asking the New Iberia centre to keep some animals for the short term. If expanded to full capacity, Chimp Haven says that it could eventually house around 430 chimps. ■

PHYSICS

Theorists bridge space-time rips

Framework offers starting point to explaining how particles cope with fluctuations in gravity.

BY EUGENIE SAMUEL REICH

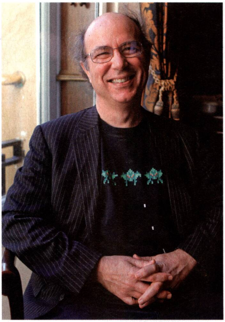
Could an analysis based on relatively simple calculations point the way to reconciling the two most successful — and stubbornly distinct — branches of modern theoretical physics? Frank Wilczek and his collaborators hope so.

The task of aligning quantum mechanics, which deals with the behaviour of fundamental particles, with Einstein's general theory of relativity, which describes gravity in terms of curved space-time, has proved an enormous challenge. One of the difficulties is that neither is adequate to describe what happens to particles when the space-time they occupy undergoes drastic changes — such as those thought to occur at the birth of a black hole. But in a paper posted to the arXiv preprint server on 15 October (A. D. Shapere *et al.* http://arxiv.org/abs/1210.3545; 2012), three theoretical physicists present a straightforward way for quantum particles to move smoothly from one kind of 'topological space' to a very different one.

The analysis does not model gravity explicitly, and so is not an attempt to formulate a theory of 'quantum gravity' that brings general relativity and quantum mechanics under one umbrella. Instead, the authors, including Nobel laureate Frank Wilczek of the Massachusetts Institute of Technology (MIT) in Cambridge, suggest that their work might provide a simplified framework for understanding the effects of gravity on quantum particles, as well as describing other situations in which the spaces that quantum particles move in can radically alter, such as in condensed-matter-physics experiments. "I'm pretty excited," says Wilczek, "We have to see how far we can push it."

The idea is attracting attention not only because of the scope of its possible applications, but because it is based on undergraduate-level mathematics. "Their paper starts with the most elementary framework," says Brian Greene, a string theorist at Columbia University in New York. "It's inspiring how far they can go with no fancy machinery."

Wilczek and his co-authors set up a hypothetical system with a single quantum particle moving along a wire that abruptly splits into two. The stripped-down scenario is effectively the one-dimensional version of an encounter with ripped space-time, which occurs when the topology of a space changes radically. The theorists concentrate on what happens at the endpoints of the wire — setting the 'boundary conditions' for the before and after states of the quantum wave associated with the particle. They then show that the wave can evolve continuously without facing any disruptions as the boundary conditions shift from one geometry to the other, incompatible one. "You can smoothly follow this process," says Al Shapere at the University of Kentucky in Lexington, a co-author on the paper, adding that, like a magician's rings, the transformation is impossible to visualize, but does make mathematical sense.


J. HERRERO/EPA/CORBIS

Frank Wilczek studies how fundamental particles respond to drastic changes in space-time.

The desire to escape the mathematical headaches caused by such transformations is one motivation for string theory, which allows smooth changes in the topology of space-time, says Greene. He suggests that the approach developed by Wilczek, Shapere and MIT undergraduate student Zhaoxi Xiong could be applied within string theory too.

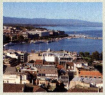
Although Wilczek originally believed that the result was new, a 1995 paper by Aiyalam Balachandran of Syracuse University in New York proposed a similar strategy for describing changes in topology in quantum mechanics (A. P. Balachandran *et al. Nucl. Phys. B* **446,** 299–314; 1995). Balachandran acknowledges that his work hasn't hit the mainstream and says that he hopes Wilczek's paper will prompt others to take a closer look. "Conventional approaches to this problem don't get very far," he says. "This opens up a new technique."

The framework might also provide inspiration for experimentalists working on condensed matter. Rob Myers, a string theorist at the Perimeter Institute for Theoretical Physics in Waterloo, Canada, says that he expects it to be relevant to an area called quantum quenches, in which quantum systems evolve in isolation from the environment and are then kicked out of equilibrium by an action of the experimentalist. Condensed-matter physicists have developed several quantum systems — including cold-atom traps and superconducting circuits — that can be used to test this idea.

Although the authors lay out their solution in only one dimension, Myers expects that the approach will readily generalize to describe real experiments in three dimensions. But he cautions that the paper represents only a first step. "To really see the impact of this work, that will take a while," he says. ■

MORE ONLINE

TOP STORY



Ancient tsunami devastated Lake Geneva shoreline go.nature.com/ptao2z

MORE NEWS

- Testing magnesium's brain-boosting effects go.nature.com/6sx71a
- ITER experiment struggles to put the pieces together go.nature.com/sgnepd
- 'Penis worm' pokes holes in evolutionary dogma go.nature.com/p3qkos

PODCAST

nature podcast

Cancer risk for redheads, mapping the bird family tree and single-cell sequencing go.nature.com/cykasc

RADIUS IMAGES/ALAMY

wellcometrust

A world where the brightest minds in science get the support they need.

Opportunities in Science Funding

The Science Funding division is the main scientific interface between the Wellcome Trust and our global biomedical and public health research community. The division, a stimulating group of 40 professional scientific administrators, is responsible for funding and supporting a large part of the biomedical science community in the UK and overseas. In 2010, we spent almost £500m supporting thousands of scientists in dozens of countries, to implement the challenges and foci of our Strategic Plan. We seek people who'll be uncompromising in pursuit of our mission, willing to question authority, and confident with their peers and the external community.

Head of Department, Science Funding Cellular, Developmental and Physiological Sciences

£75 000+ per annum

In this key role, reporting to the Director of Science Funding, you'll manage a departmental staff to develop our portfolio in cellular, developmental and physiological sciences – seeking out excellence, catalysing opportunities and supporting our grant holders in these scientific areas. The role is based at our landmark Gibbs Building headquarters in central London, but will involve some UK travel. To meet the challenge, you'll need a doctoral degree and significant research experience, as well as experience in science administration – possibly leading your own research group.

Science Portfolio Manager, Science Funding Deputy Head of Department, Basic Science Careers

£50 000+ per annum

Working closely with our Head of Basic Science Careers, you'll develop and support our portfolio of training and early career schemes in the basic sciences, including PhD, post doctoral and intermediate fellowship programmes. You'll need a doctoral degree and significant research experience, in addition to excellent interdisciplinary knowledge. You'll also need outstanding interpersonal, organisational and analytical skills and judgement, and direct experience of supervising research students and/or post docs.

Find out more and apply online at **www.wellcome.ac.uk/jobs**.

Alternatively, send a CV and covering letter, stating clearly the role for which you'd like to be considered and explaining how you meet the criteria, to HR, Wellcome Trust, 215 Euston Road, London NW1 2BE or to **HR@wellcome.ac.uk**.

Closing date: **Friday 16 November 2012**

A. OHANESIAN/REUTERS

Poor sanitation at the Yusuf Batil refugee camp in South Sudan caused an outbreak of hepatitis E earlier this year.

IMMUNOLOGY

Hepatitis E vaccine debuts

Success of Chinese biotech partnership raises hopes for prevention of overlooked diseases.

BY SOO BIN PARK

Batches of the world's first vaccine against the hepatitis E virus began rolling out of a Chinese factory last week, promising to stem a disease that every year infects about 20 million people and claims 70,000 lives. The vaccine is being hailed as a victory for an unusual public–private partnership that could set a precedent in China's burgeoning biotechnology sector, and help to deliver other vaccines for diseases overlooked in the West.

The waterborne hepatitis E virus mostly occurs in developing countries that have poor sanitation, and it is particularly prevalent in east and south Asia. Although most cases cause only mild illness, it can lead to acute liver failure — the mortality rate reaches 4% in some regions and soars to 20% in women who are in the later stages of pregnancy. A severe outbreak of hepatitis E in the Xinjiang Uygur Autonomous Region in northwest China[1], for example, caused almost 120,000 infections and more than 700 deaths between 1986 and 1988 (see 'Hidden epidemics'). There is no treatment, and improved sanitation has so far been the most effective way to stem the disease.

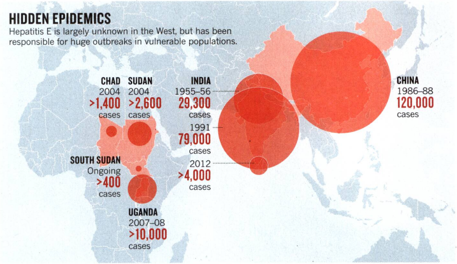
The new vaccine, which was approved by China's State Food and Drug Administration (SFDA) in December 2011, could transform that picture. More than a decade ago, researchers at Xiamen University in Fujian province genetically modified a strain of the bacterium *Escherichia coli* to produce a protein that, when injected into humans, stimulates the body's immune system against hepatitis E. But preclinical and clinical development began in earnest only in 2000, when the Yangshengtang Group, a company with interests in food and health care, invested 15 million renminbi (US$1.8 million in 2000) to set up a joint biotech laboratory in partnership with the university. The lab was given national status in 2006 by the Chinese Ministry of Science and Technology and relaunched as the National Institute of Diagnostics and Vaccine Development in Infectious Diseases (NIDVD).

The institute aims to unite academia and industry in commercializing new vaccines, particularly for emerging infectious diseases. Yangshengtang set up a subsidiary company called Innovax to take potential vaccines through clinical trials to manufacturing. The hepatitis E vaccine, Hecolin, is the company's first product to reach the market, but it also has a vaccine against human papilloma virus that is currently in preclinical research. Approval for Hecolin came after a phase III clinical trial published in 2010 showed that it was highly effective in preventing infection among almost 100,000 healthy participants[2].

Hecolin cost about 500 million renminbi (US$80 million) to develop, much of which came from the Chinese government through the university. The vaccine will be sold to distributors in China at a cost of 110 renminbi per dose, and the company expects it to reach sales of 62 million renminbi in 2013. That is hardly a blockbuster income, but, according to Jun Zhang, deputy director of the NIDVD, the public–private development model helps to ensure that vital vaccines are developed regardless of whether they prove to be profitable for manufacturers.

Zhang hopes that the success of Hecolin will attract further investment in such schemes, and says that the Chinese government has been encouraging. "Many people — including representatives of multinational pharmaceutical companies, venture capitalists, Chinese local government officials and Chinese entrepreneurs — think this is a worthy example of biotechnology investment," he says.

Zhang points out that UK drug-maker ▶


HIDDEN EPIDEMICS

Hepatitis E is largely unknown in the West, but has been responsible for huge outbreaks in vulnerable populations.

SOURCES: J. M. HUGHES *ET AL. CLIN. INFECT. DIS.* **51,** 328–334 (2010)/WHO/PROMED-MAIL

▶ GlaxoSmithKline had already developed a separate hepatitis E vaccine in collaboration with the US Army, which showed promise in phase II trials[3]. But with hepatitis E mostly occurring in developing countries, there was little commercial potential for the vaccine. "This is true not just of hepatitis E, but also many other plagues in the world," says Zhang.

Medical products for conditions such as hepatitis E that predominantly affect the developing world "are not seen as big money opportunities", agrees Jeremy Farrar, director of the Oxford University Clinical Research Unit in Ho Chi Minh City, Vietnam. "New companies operating with different funding models offer a great opportunity, and one which could have a profound impact."

Hecolin may have arrived just in time to tackle a rise in hepatitis E in Africa, where a 2007 outbreak[4] in Uganda infected more than 10,000 people and killed 160. By the end of September this year, more than 200 cases of jaundice caused by hepatitis E had been reported in refugee camps in Kenya since August, and three refugee camps in South Sudan have seen 16 deaths and 400 cases of hepatitis E infection since July. "Cases are rising day by day, thus placing immense pressure on the available health services and resources. This is of grave humanitarian concern," said South Sudan's health ministry in a statement in September.

Xiamen University and Innovax are in talks with the World Health Organization (WHO) to register Hecolin with the organization's Prequalification Programme, which makes medicines available to agencies such as the United Nations Children's Fund and the Joint UN Programme on HIV/AIDS. "We have to be sure that these vaccines can be used anywhere," says Farrar. "It would be a great shame if these products were not available outside China."

"We have to accept that companies such as this one in China are going to be very important in the future," he adds. "The rest of us have to catch up. We need to find a way, through the WHO, of ensuring the absolute transparency, safety and effectiveness of their vaccines." ■

1. Zhuang, H., Cao, X. Y., Liu, C. B. & Wang, G. M. *Gastroenterol. Jpn.* **26,** 135–138 (1991).
2. Zhu, F.-C. *et al. Lancet* **376,** 895–902 (2010).
3. Shrestha, M. P. *et al. N. Engl. J. Med.* **356,** 895–903 (2007).
4. Teshale, E. H. *et al. Emerg. Infect. Dis.* **16,** 126–129 (2010).

FOOD SAFETY

Bid to curb fried-food chemical goes cold

Acrylamide levels still too high in Europe's food, says report.

BY KATHARINE SANDERSON

The rich, roasted aroma of coffee or the golden-brown colour of crispy French fries are enough to set most mouths watering. But the high-temperature cooking that gives these foods their alluring taste, scent and texture also adds a sting: acrylamide, a probable human carcinogen.

Swedish scientists discovered in 2002 that a wide range of baked and fried goods contain worryingly high levels of acrylamide[1] — a simple organic molecule that is a neurotoxin and carcinogen in rats. The finding sparked an international effort to reduce concentrations of the chemical by changing ingredients and cooking methods.

Ten years on, a report[2] from the European Food Safety Authority (EFSA) in Parma, Italy, suggests that this effort has stalled, amid patchy monitoring, uncertainty about acrylamide's true health effects and the challenge of weeding out a molecule present in hundreds of products.

Soon after the Swedish discovery, two teams — one led by chemist Donald Mottram at the University of Reading, UK, the other by Richard Stadler at Nestlé in Lausanne, Switzerland — unpicked the chemistry behind the problem[3,4]. They found that sugars and amino acids such as asparagine found in potatoes and cereals were making acrylamide (C_3H_5NO) as a by-product of the Maillard reaction, the very process that generates the heady blend of colour, flavour and taste in cooked foods.

Subsequent epidemiological studies involving tens of thousands of people have looked for links between acrylamide and various forms of cancer in humans, including breast[5] and colorectal cancer[6]. For the most part, the results have been negative. In 2007, however, a Dutch study[7] of almost 2,600 women found that, among those who had never smoked, women consuming about 40 micrograms of acrylamide per day doubled their risk of developing cancers of the womb or ovaries, compared with those taking in roughly 10 μg per day. And last month, a study[8] showed that women who ate acrylamide-rich food during pregnancy tended to give birth to smaller babies.

Despite the uncertainties over the dangers of acrylamide, Europe's legislators and food producers vowed to take action. Since 2005, the industry group FoodDrinkEurope has maintained a 'toolbox' of tactics to help reduce acrylamide levels, such as changing potato varieties or storage conditions, and reducing cooking temperatures. According to Beate Kettlitz, the group's director of food policy, 90% of large and medium-sized companies in Europe now select potato varieties with low levels of the sugars that can form acrylamide, and all control French-fry cooking times to limit browning.

In 2007, the European Commission instructed the EFSA to collate yearly data on acrylamide levels. Last week, the authority released the most recent figures[2] showing that acrylamide levels in finished food products hardly changed between 2007 and 2010. There have been isolated successes: in soft bread, for example, mean acrylamide levels dropped from 75 to 30 μg per kilogram. But for crispbreads, the mean actually rose, from 232 to 249 μg kg^{-1}. Overall, 6–17% of the food categories tested exceeded 'indicative values of concern'[9] set out by the European Commission in 2011 (see 'Would you like acrylamide with that?').

Mottram, who has worked closely with the food industry to reduce acrylamide levels, says that he is disappointed the report does not reflect the huge strides taken by industry, not least in the period 2002–06.

The EFSA acknowledges that assessing whether the industry's efforts are bearing fruit will take many years of more consistent sampling. For the current report, the agency relied on European Union (EU) member states to collect and submit acrylamide data. The response was inconsistent: only 16 of 25 countries provided data for every year of the survey, and submissions also waned over time. Despite submitting data every year, Belgium provided no figures at all on its beloved frites, for example.

Since 2010, the EU has required member countries to collect acrylamide data, and the EFSA report suggests that monitoring is improving as a result. Europe certainly takes acrylamide more seriously than other parts of the world. The US Food and Drug Administration has not routinely collected data on acrylamide in food since 2006, although it is currently calling on the food industry to submit more data, says agency spokesman Sebastian Cianci.

WOULD YOU LIKE ACRYLAMIDE WITH THAT?

Figures for 2007–10 suggest that fried and baked foods in Europe often contain worryingly high levels of the probable carcinogen acrylamide.

Food	Indicative value of concern (μg kg^{-1})	% Samples exceeding indicative value
French fries	600	12
Potato crisps	1,000	17
Instant coffee	900	10
Soft bread	150	7
Crispbreads, biscuits	500	8

Source: ref. 2

Mottram is also developing new ways to tackle the problem. In August, he showed that acrylamide levels in French fries can be predicted from the cooking methods and the presence of key precursor chemicals in the partially cooked, frozen fries used by fast-food restaurants[10]. This model revealed that a change to the potato blanching process could make a big difference to the final acrylamide level. "The industry is not giving up on this," he says.

Plant breeding and genetic modification could also help, by creating varieties with lower levels of acrylamide precursors. Nevertheless, the chemical will always be present in our food, says Margareta Törnqvist from Stockholm University, who led the team that originally discovered the problem. "Acrylamide is natural; you can't reduce it to zero," she says. ■

1. Tareke, E., Rydberg, P., Karlsson, P., Eriksson, S. & Törnqvist, M. *J. Agric. Food Chem.* **50,** 4998–5006 (2002).
2. European Food Safety Authority *EFSA J.* **10,** 2938 (2012).
3. Mottram, D. S., Wedzicha, B. L. & Dodson, A. T. *Nature* **419,** 448–449 (2002).
4. Stadler, R. H. *et al. Nature* **419,** 449–450 (2002).
5. Wilson, K. M. *et al. Am. J. Epidemiol.* **169,** 954–961 (2009).
6. Larsson, S. C., Åkesson, A., Bergkvist, L. & Wolk, A. *Eur. J. Cancer* **45,** 513–516 (2009).
7. Hogervorst, J. G., Schouten, L. J., Konings, E. J., Goldbohm, R. A. & van den Brandt, P. A. *Cancer Epidemiol. Biomarkers Prev.* **16,** 2304–2313 (2007).
8. Pedersen, M. *et al. Environ. Health Perspect.* http://dx.doi.org/10.1289/ehp.1205327 (2012).
9. *Commission Recommendation of 10.1.2011 on Investigations into the Levels of Acrylamide in Food* (European Commission, 2011); available at go.nature.com/tkdm89.
10. Parker, J. K. *et al. J. Agric. Food Chem.* **60,** 9321–9331 (2012).

CORRECTION

The story 'Texas cancer fund seeks fresh start' (*Nature* **490,** 459–460; 2012) gave the wrong location for Arizona State University with respect to Raymond DuBois' new post. The campus is in Tempe, not Tucson.

High impact science
for the Arabic community

nature
الطبعة العربية

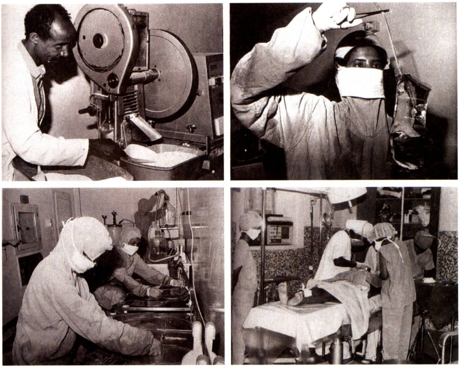
Nature Arabic Edition allows Arabic speakers throughout the world to access top quality science news and comment from *Nature*, as well as summaries of all the research papers from the leading multidisciplinary journal. Content is freely available online every week, with limited monthly copies available in print.

Access *Nature Arabic Edition* online and apply for your free subscription **arabicedition.nature.com**

In partnership with:

مدينة الملك عبدالعزيز
KACST للعلوم والتقنية

nature publishing group npg



High-quality pharmaceuticals and intravenous fluids were prepared and surgeries performed in Eritrea's 'longest hospital in the world' during the country's fight for independence.

ANGESOM TEKLEHAIMANOT BOKRU

ERITREA'S SHATTERED SCIENCE

BY SHANTA BARLEY

An impoverished African nation was making promising strides in medicine — before the government clamped down on its foreign partnerships.

Early this year, Eritrea severed a scientific lifeline almost as old as the African nation itself. The Eritrean National Health Laboratory in Asmara cut long-standing ties with Washington University School of Medicine in St Louis, Missouri, potentially setting back many gains that the country had made in public health. "St Louis supplied everything: American doctors, expertise, chemicals, materials," says Assefaw Ghebrekidan, an Eritrean ex-freedom fighter who now heads the public-health programme at Touro University in Mare Island, California. "And now it's all over."

Eritrea, an impoverished country of 3 million people on the Horn of Africa (see 'A troubled corner'), is not known for its science. It ranks 177th out of 187 countries on the United Nations Human Development Index. It comes in last in terms of press freedom and is the eighth most militarized country in the world. The World Health Organization estimated that there were just 5 medical doctors per 100,000 people in the country in 2004.

But against this depressing backdrop, the country's medical-research partnerships have been a source of promise and pride. Eritrea built its first medical school in 2003, aided by scientists from the Central University of Las Villas in Santa Clara, Cuba. After US universities helped to establish postgraduate training and research programmes in paediatrics, surgery, and obstetrics and gynaecology at the institution, Eritrean medical scientists published their first papers in international, peer-reviewed journals. Public health has benefitted. In 1991, Eritrea was cursed with the highest maternal mortality rate in the world — 14 deaths per 1,000 births. In 2010, it was on track to meet the Millennium Development Goal of cutting that rate by 75% by 2015.

But progress in Eritrean science has now gone into reverse, say a number of scientists and doctors in exile. In response to mounting criticism from the United Nations and the United States over the country's human-rights record, Eritrean President Isaias Afwerki is severing partnerships with all US universities, says Ghebrekidan. "Everything that Eritrea has worked so hard to achieve is at stake."

Jon Abbink, an anthropologist at the Free University of Amsterdam, says that these actions will have widespread negative effects, "in the education system, in the constant 'brain drain' of educated people to greener and freer pastures, and in the inhibition of international scientific cooperation". Eritrea, he says, is one of the few remaining countries in Africa that have failed to embrace scientific freedom. "It's out of sync with global trends," says Abbink.

Eritrea was once a colony of Italy, but the United Nations handed it over to Ethiopia after the Second World War. In 1961, Eritrea started to fight for its independence in a war that would last three decades: the United States supplied Ethiopia with guns and money, but the rebels, led by Afwerki and the Eritrean People's Liberation Front (EPLF), persevered.

The liberation movement had remarkable credentials. "It was led by 29 doctors of medicine," says Ghebrekidan, who was head of the EPLF's medical services. "No other rebel movement has ever had so many intellectuals." Even Afwerki had abandoned a degree in engineering to lead the fight.

Another academic, Melles Seyoum, was working as a pharmacist at an Ethiopian hospital when the war broke out. He coolly stole US$140,000 worth of antibiotics, microscopes, surgical blades and stethoscopes and delivered them to Eritrean freedom fighters, wrote journalist Michela Wrong in her book *I Didn't Do It For You* (HarperCollins, 2005). Seyoum became an integral member of the EPLF, teaching soldiers how to test blood and prepare Petri dishes in a hospital 5 kilometres long and dug into the side of a rocky valley — a clinic known as 'the longest hospital in the world'. After a visit in 1987, a British doctor wrote[1] about the impressive standards of care at the hospital: a 1-tonne machine manufactured antibiotics every day; a doctor performed facial reconstructions; and amputees played basketball.

POWER STRUGGLE

In 1993, after the war ended and Eritrea gained independence, Afwerki was elected president by a national assembly largely composed of former members of his rebel army. He promised that within four years, Eritrea would have parliamentary and presidential elections, press laws and a new constitution. Seyoum enthusiastically backed Afwerki and was, in return, appointed director of the prestigious National Health Laboratory, which performed most of the country's clinical testing and worked on developing treatments for disease.

But following a failed assassination attempt in 1996, the president postponed elections indefinitely and refused to implement the constitution that had been drafted. In 1998, he invaded Ethiopia, triggering a humiliating two-year war that caused the deaths of more than 60,000 Eritreans and a temporary loss of one-quarter of Eritrean territory. Afwerki's popularity plummeted, and many of the academics who had helped to rebuild the country moved abroad.

On 3 October 2000, some of them decided to use their friendship with Afwerki to persuade him to hold elections or step down. From a conference hotel in Berlin, Ghebrekidan and 12 other scientists and professionals, many of whom had been involved in drafting the constitution, composed a letter to the president.

"Much of the world community, including our fellow Africans, perceive the Eritrean government and its leadership as aggressive and irresponsible," wrote the group, urging Afwerki to implement the constitution, hold democratic elections and set free the growing number of people his regime had jailed. "We urge you most sincerely to seize this moment of crisis and turn it into an opportunity to reclaim your hard-earned reputation as a leader." Four days later, after it had reached Afwerki, the letter was leaked to the press, igniting Eritrea's first-ever public debate about leadership.

To its members' surprise, the group — which became known as the G-13 — was invited to Eritrea for discussions with Afwerki. One member, Mohammed Kheir, later wrote that he was nervous that it might be a trap. But they accepted the invitation and flew to Eritrea. After waiting for several days, the president agreed to see them. Soldiers escorted the academics to his office, where Afwerki berated them for leaking the letter to the media — something that they denied — and cast them as traitors. The group was escorted back to the airport. Since then, no members have returned to Eritrea; most now hold prestigious positions at US universities. "It is very fortunate that we escaped," says Haile Debas, now head of the University of California Global Health Institute in San Francisco.

Although the plea failed to sway the president, it encouraged others to criticize him openly for the first time. In July 2001, Semere Kesete, leader of the student union at the University of Asmara — Eritrea's only institute of higher learning — criticized the government for reducing academic freedom. He was arrested and thrown into solitary confinement, causing riots at the university. When the government demanded that the students do extra national service — on top of the 18 months required of all men and women — they didn't turn up. In retaliation, the government bussed all of the students to the Danakil Depression in southern Eritrea, one of the hottest places on Earth, to build roads. Two students died from the heat.

CRACKDOWN

A month later, Afwerki launched his biggest crackdown yet. He shut down all private media, threw 10 journalists in jail and imprisoned 11 politicians who had demanded elections — many of whom were old comrades in arms. He also began to dismantle the University of Asmara.

"What could be the justification for killing the only university we had capable of producing students that could be accepted by universities abroad?" asks an Eritrean scientist who lives out of the country and wishes to remain anonymous because of concerns about the safety of family members still in Eritrea. "The aim was simply to prevent the students from all being in one place, where they had the power to rise up," says Debas. In place of the university, the government built a number of small colleges, arguing that these would be more accessible to students.

Even as Eritrea lost its only university, it continued to make progress in medicine. In 1997, the country had gained a proactive health minister, Saleh Meki, who helped to develop crucial partnerships with US universities including George Washington University in Washington DC; Washington University in St Louis; Columbia University in New York City; Stony Brook University in New York; and the University of California, Berkeley. By bringing experts into Eritrea, these partnerships helped the country to pass scientific and public-health milestones. A polio-immunization campaign extended coverage to 95% of all one-year-olds and eradicated the disease. An anti-malaria drive from 2000 to 2004 reduced morbidity and case fatality by 84% and 40%, respectively.

SUDAN
Red Sea
ERITREA
Asmara
ETHIOPIA

A TROUBLED CORNER

Eritrea won independence from Ethiopia in 1993 after 30 years of war.

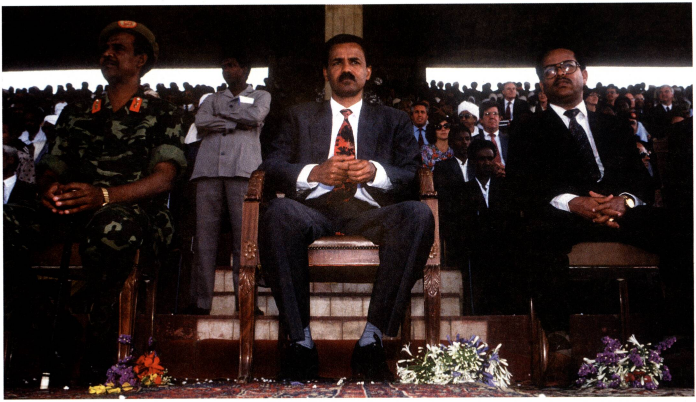
In 2003, Haile Mezgebe, then a surgeon at George Washington University, was part of the group of medics who helped to set up the Orotta School of Medicine in Eritrea. Mezgebe moved to the country to run the collaboration; he was joined by Mary Polan, who travelled regularly from the department of obstetrics and gynaecology at Columbia University, and other US doctors and surgeons who worked to treat and train Eritreans. In 2009, Orotta graduated its first class of 39 doctors. "It was quite extraordinary," says Jack Ladenson, a doctor based at Washington University. "Suddenly, in one day, there was a 30% increase in the number of doctors in Eritrea."

SUCCESS STORY

Meanwhile, clinical testing and research was taking off at the National Health Laboratory. In 1998, the only blood tests available in Eritrea had been done on a single machine. Scientists from Washington University installed new equipment at the lab and trained technicians to perform a range of chemical tests, including the haemoglobin A_1C test for diabetes and a test for thyroid malfunction. They also launched a national diabetes-management programme and a long-term research project to gauge its progress; in 2007, the project leaders found[2] that the programme had significantly improved Eritrean diabetes management. Ladenson, Seyoum and others co-authored a paper[3] showing that the overall quality of chemical tests for disease at the national lab was on a par with that at Washington University. "A simple but sustainable national laboratory system has been established in the developing nation of Eritrea," the paper said.

But outside the medical arena, the situation was less rosy. Richard Reid, a historian at the School of Oriental and African Studies in London, visited the Eritrea Institute of Technology in Mai Nefhi, one of the unaccredited colleges set up after the University of Asmara was shut down. He was told that students who cheated on exams or skipped classes were jailed on site. Military training was mandatory between 4 and 7 a.m., and students wryly referred to digging trenches as 'digology', adds Reid.

And any success in science and medicine was short lived. In 2008, without explanation, Meki was removed as health minister, along with the coordinator for US–Eritrean scientific



S. PETERSON/LIAISON/GETTY

Isaias Afwerki (centre) in 1992, a year before he became Eritrea's first president. He has recently halted the country's collaborations with US medical schools.

partnerships. The chair of the paediatrics department at Orotta was arrested because of his religious views. And in 2011, Afwerki ordered all scientists from George Washington University — including Mezgebe — to leave the country.

> "THE GOVERNMENT HAS PERSECUTED NOT ONLY SCIENTISTS, BUT ALSO THE SCIENCE."

At the start of 2012, Afwerki cut off the partnership between the National Laboratory and Washington University in St Louis. Several sources, who wish to remain anonymous for fear of retaliation against friends and relatives, report that Seyoum, the lab's director, was "frozen", an Eritrean term for the practice of stripping government employees of their titles and duties while restricting them from travel and other jobs to silence them. *Nature* contacted officials in the Eritrean government and its US and UK embassies repeatedly by phone and e-mail for a response to these allegations, but had received none at the time of going to press.

The severing of ties may be a backlash against the United States and the United Nations over their criticism of Afwerki's human-rights record, says Ghebrekidan. In 2009, the United States imposed sanctions on Eritrea for supporting Islamist insurgents in Somalia. A highly publicized cable from US ambassador Ronald McMullen, later released by Wikileaks, said that "Eritrea's prisons are overflowing, and the country's unhinged dictator remains cruel and defiant". In July, the UN Human Rights Council established a special rapporteur to investigate reports of rights violations by Eritrean authorities, amid stories that Afwerki keeps his critics in solitary confinement in shipping containers.

Berhane Ghebrehiwet, an Eritrean immunologist at Stony Brook University, says that Afwerki's distrust of foreign involvement and aid in Eritrea is understandable. The United States did, after all, support Ethiopia during the fight for independence. "You cannot cripple a man and then accuse him of having limped," he says. "All the president dreams of is to make Eritrea a prosperous and self-reliant nation at peace with itself, its neighbours and the rest of the world."

Others are less sympathetic. "Afwerki is getting more and more paranoid," says Ghebrekidan. "He thinks that the American doctors who come to save Eritrean lives are actually CIA agents."

Afwerki has effectively destroyed intellectual freedom in Eritrea, says Abbink. "No independent academic research in any field is possible." Fundamental research "or what is left of it" is now under pressure to pursue "practical" issues with immediate applications to development, he concludes.

Yet some scientists are still proud of the progress Eritrea has made. Andemariam Gebremichael, dean of the Orotta School of Medicine, wrote in an e-mail that he aims to create "an environment where individuals develop their intellectual potential", adding that he hopes to produce another 150 doctors to bring the country up to international standards. It will be a significant challenge, writes Gebremichael. Only seven foreign teaching doctors — all Cuban — remain at the institution.

After a year in solitary confinement, student-union leader Kesete escaped with his guard to Ethiopia, and from there to the United States. "We walked for six days and nights, surviving on nothing more than biscuits," he says.

Kesete sees little prospect of change, and despairs of his country's future. "The government has persecuted not only scientists, but also the science itself," he says. He calls international collaborations a "waste of resources and energy", because Afwerki will not hesitate to eject foreign scientists, no matter how crucial they are to Eritrea's development. Rumours that the University of Asmara may reopen this year are preposterous, he adds. "It is safe to say that academia is dead." ■

Shanta Barley *is a freelance writer in Perth, Australia.* SEE EDITORIAL P.8

1. Powell, H. D. W. *Br. Med. J.* **295,** 1637–1639 (1987).
2. Windus, D. W. *et al. Clin. Chem.* **53,** 1954–1959 (2007).
3. Scott, M. G., Morin, S., Hock, K. G., Seyoum, M. & Ladenson, J. H. *Clin. Chem.* **53,** 1945–1953 (2007).

VISUALS UNLIMITED/CORBIS

THE SINGLE LIFE

Sequencing DNA from individual cells is changing the way that researchers think of humans as a whole.

BY BRIAN OWENS

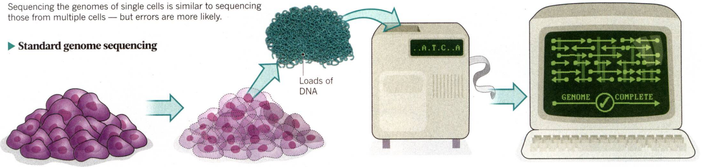
All Nicholas Navin needed was one cell — the issue was how to get it. It was 2010, and the postdoctoral fellow at Cold Spring Harbor Laboratory in New York was exploring the genetic changes that drive breast cancer. Most of the cancer-genome studies before then had ground up bits of tumour tissue and sequenced the DNA en masse, giving a consensus picture of the cancer's genome. But Navin wanted to work out the sequence from individual cells to see how they had mutated and diverged as the cancer grew.

He ran into trouble almost immediately. "Cells like to stick together," he says. He tried the most advanced microdissection techniques, which use robots to peel cells apart or suck them into the tips of hair-thin glass pipettes. But he could never be sure that a second cell hadn't come along for the ride. Eventually, he settled on using chemicals to dissolve the cell's outer membrane and release the dense nucleus. Then he separated out the nucleus

ONE GENOME FROM MANY

Sequencing the genomes of single cells is similar to sequencing those from multiple cells — but errors are more likely.

▶ Standard genome sequencing



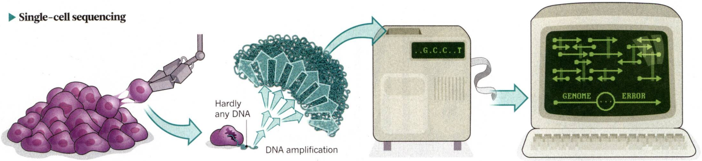
A sample containing thousands to millions of cells is isolated.

DNA is extracted from all the nuclei.

DNA is broken into fragments and then sequenced.

The sequences are assembled to give a common, 'consensus' sequence.

▶ Single-cell sequencing



A single cell is difficult to isolate, but it can be done mechanically or with an automated cell sorter.

The DNA is extracted and amplified, during which errors can creep in.

Amplified DNA is sequenced.

Errors introduced in earlier steps make sequence assembly difficult; the final sequence can have gaps.

using an automated cell sorter, and extracted its DNA. He repeated the process for around 100 cells, and the sequences he obtained revealed how the tumour had evolved from a few rogue cells into a complex mélange of genetically distinct ones[1].

The ability to sequence 100 human cancer genomes was unthinkable a decade ago, and it is still a remarkable feat. Technology has moved apace, dramatically reducing costs and making genome sequencing fairly routine. But most human genomes, cancer or otherwise, are still sequenced from DNA extracted from multiple cells, which misses differences between cells that could be crucial in controlling gene expression, cell behaviour and drug response.

"I THINK IT'S THE NEXT LEVEL OF COMPLEXITY."

"People are becoming very interested in what is the variation from cell to cell," says Navin, now at the University of Texas MD Anderson Cancer Center in Houston. Last month, the US National Institutes of Health (NIH) announced that it would be funding single-cell studies, including genome sequencing, to the tune of US$90 million over 5 years.

But challenges abound. Amplifying the tiny amount of DNA in a single cell until there is enough to sequence without introducing too many errors is still difficult (see 'One genome from many'). The bioinformatics required to stitch the data together and deal with artefacts can be fiendishly complicated. And, as Navin found, even isolating a cell can be tough. For this reason, several research groups have started with cells that are easily separated, such as sperm, or those that are likely to have dramatic genomic differences, such as tumour cells.

But as the techniques are refined, scientists hope to work out ever more subtle differences between cells, such as the tiny genomic rearrangements that happen in many neurons and may serve a purpose in organizing information flow. One of the most startling facets of such research is not the differences between cells, but how tissues and organs manage to work coherently despite them. To get a handle on these issues, it makes sense to start at the smallest possible unit. "The cell is kind of the ultimate denomination of an organism," Navin says.

TUMOUR DIVERSITY

In his first effort with breast-cancer tumours[1], Navin was able to sequence only about 10% of the DNA — not enough to see individual point mutations, but good enough to study larger segments that are commonly duplicated or deleted, called copy number variants.

The results suggested that the tumour was made up of three major populations of cells, which emerged from the root tumour population in leaps and spurts at different times during the tumour's growth. "It suggested a model of evolution where, instead of having lots of gradual intermediates, we saw hundreds of chromosomal rearrangements that occurred probably in very short periods of evolutionary time," he says.

Since moving to Texas, Navin has started a research group focused entirely on single-cell genomics. He has been improving his methods, and can now piece together up to 90% of a cell's genome, he says, which allows him to study the mutations in individual cancer cells in much more detail.

The team has also looked at another type of breast cancer. Sequencing the tumour as a whole, the group found six mutations in cancer-associated genes. "It seemed like a very simple genome," says Navin. But when his group sequenced four individual cells they identified "hundreds of additional mutations, many of which where unique or 'private' to that individual cell". In addition to tumours, his group has studied a breast-cancer cell line that has been maintained in the laboratory. The cells in such lines might be expected to contain identical genomes, but in data that have not yet been published, the team found that about 1% of the mutations, involving 12,000–20,000 base pairs, differed from cell to cell and that these variations could not be detected when the cells were sequenced together. In both studies, many of the new-found mutations were in cancer genes, or in other regions expected to disrupt protein function.

"Nicholas's work was very nice," says Timour Baslan, a graduate student studying cancer genetics at Cold Spring Harbour Laboratory and a former colleague of Navin's in Michael Wigler's lab. "It was very informative, but it was extremely expensive" — $1,000 or more per cell. Baslan and his colleagues are trying to bring down the cost. They add genetic barcodes — short, easily traceable strands of DNA — to a cell's DNA, which allows them to sequence cell genomes en masse then identify the sequences from individual cells. Use of the barcodes, in combination with improved bioinformatics approaches, brings the cost of a genome down to about $60 per cell. At that price, more researchers can work with the number of cells necessary to dissect a tumour-cell population as it proliferates and evolves. "Now we're doing hundreds and we're thinking in the future we'll be able to do thousands," says Baslan.

Baslan's group is using the techniques to study which tumour cells are left behind after chemotherapy treatment and why those cells were resistant to the drugs. Such analysis could guide treatment, says Navin. "By sequencing a few single cells you can get an idea of the heterogeneity in a tumour before chemotherapy and that might affect your choice of which agent to use, or whether to subject the patient to chemotherapy at all," he says.

EVERY SPERM IS SACRED

The tendency of sperm to swim alone makes the cells ideal for single-cell genomics. Adam Auton, a statistical geneticist at Albert Einstein College of Medicine in New York is using sperm to study recombination, the process that shuffles genes during the formation of germ cells and therefore influences which genes are inherited. "Recombination is one of the fundamental forces that shapes genetic diversity," he says. "In recent years we've learned that there is considerable variation in the recombination rate between different populations, between the sexes and even between individuals." But pinning down the rate in people once seemed impossible because it would have required finding individuals with hundreds of children and sequencing their genomes.

The ability to sequence single cells meant that researchers could take another approach. Working with a team at the Chinese sequencing powerhouse BGI, Auton sequenced nearly 200 sperm cells and was able to estimate the recombination rate for the man who had donated them. The work is not yet published, but Auton says that the group found an average of 24.5 recombination events per sperm cell, which is in line with estimates from indirect experiments[2]. Stephen Quake, a bioengineer at Stanford University in California, has performed similar experiments in 100 sperm cells and identified several places in the genome in which recombination is more likely to occur. The location of these recombination 'hotspots' could help population biologists to map the position of genetic variants associated with disease.

Quake also sequenced half a dozen of those 100 sperm in greater depth, and was able to determine the rate at which new mutations arise: about 30 mutations per billion bases per generation[3], which is slightly higher than what others have found. "It's basically the population biology of a sperm sample," Quake says, and it will allow researchers to study meiosis and recombination in greater detail.

Perhaps the most intriguing potential use of single-cell sequencing lies in neuroscience. Alysson Muotri, a neuroscientist at the University of California, San Diego, would like to study how long interspersed nuclear elements (LINEs) — 'jumping' genes that can move around the genome — cause each neuron to differ from its neighbours. His group has compared the number of LINEs in human brain, heart and liver tissue, and found that brain tissue contains significantly more jumping genes than the others[4]. Each human neuron probably has between 80 and 300 unique insertions, he says, differences that could affect a person's susceptibility to neurological disorders[5], or provide the brain with a reservoir of diversity with which to respond to challenges. He says that it would be useful to sequence individual neurons and work out what effect this heterogeneity is having on brain function and even on personality. "I think it's the next level of complexity," he says. "We look at the brain and we think about the tissue, but actually it seems like lots of tissues in one, because the cells are so heterogeneous. It's almost like every cell was there for a purpose."

"EVERY TIME WE LOOK AT OUR DATA WE DISCOVER SOMETHING NEW."

To make good on these plans, however, the techniques will need to improve further. Ramunas Stepanauskas, director of the Single Cell Genomics Center at the Bigelow Laboratory for Ocean Sciences in East Boothbay, Maine, says that the processes involved will continue to be integrated and miniaturized, until eventually they will form an off-the-shelf technology that involves just the press of a button — although he cautions that this is still many years off. In the meantime, his centre offers single-cell sequencing for labs that lack the necessary equipment and expertise to do it themselves. The US Department of Energy's Joint Genome Institute in Walnut Creek, California, offers a similar service.

With its new funding for single-cell studies, the NIH is attempting to spur innovation in the field. So is Life Technologies, a sequencing company based in Carlsbad, California, which is offering a $1 million prize to the first researchers who can sequence the whole genome and all the RNA in a single human cancer cell using the company's technology. The deadline for the challenge is the end of this year, and the company says that there has been an enthusiastic response.

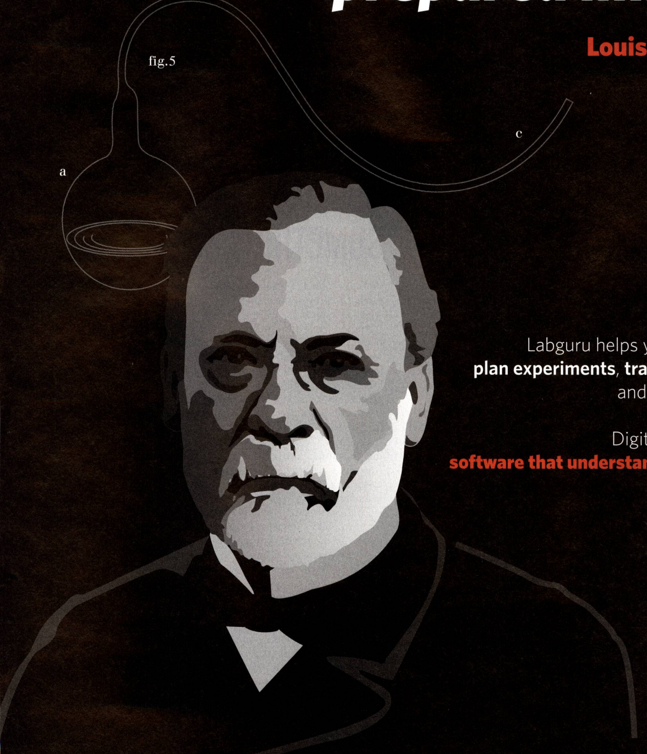
Baslan, for one, is intent on breaking down tissues into their components. "Every time we look at our data we discover something new," he says. "There's just so much analysis to be done that it's quite daunting." ■

Brian Owens *is the assistant news editor for* Nature *in London.*

1. Navin, N. *et al. Nature* **472,** 90–94 (2011).
2. McVean, G. A. T. *et al. Science* **304,** 581–584 (2004).
3. Wang, J., Fan, H. C., Behr, B. & Quake S. R. *Cell* **150,** 402–412 (2012).
4. Coufal, N. G. *et al. Nature* **460,** 1127–1131 (2009).
5. Muotri, A. R. *et al. Nature* **468,** 443–446 (2010).

"Chance favours the **prepared mind**"

Louis Pasteur



Labguru helps you to easily **plan experiments**, **track progress** and **get results**.

Digital Science – **software that understands science**.

DIGITAL science

www.digital-science.com

labguru do more science

www.labguru.com

COMMENT

POLITICS US abortion lobby set to clash with prenatal diagnostics **p.33**

DEVELOPMENT Science must inform UN sustainability goals **p.35**

NEUROSCIENCE How does the mind extract meaning from language? **p.36**

HISTORY A masterful take on the age-old allure of alchemy **p.38**

ABORTION
IS
HOMICIDE

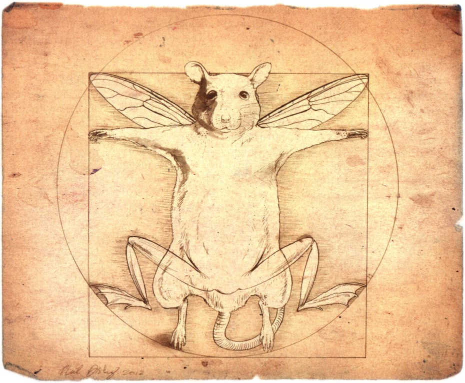


ILLUSTRATION BY PHIL DISLEY

There's more to life than rats and flies

The tiny number of model organisms constrains research in ways that must be acknowledged and addressed, warns **Jessica Bolker**.

For most experimental biologists, life revolves around a handful of species: the mouse (*Mus musculus*), the nematode worm (*Caenorhabditis elegans*), the fruitfly (*Drosophila melanogaster*) and the thale cress (*Arabidopsis thaliana*). We assume that model organisms offer universal insights, and funding agencies largely support work on a shortlist of favoured species (www.nih.gov/science/models).

Scientists who submit grant proposals for a project using a standard model organism need not use up space to explain their choice. By contrast, choosing a less common model that is uniquely suited to the research demands a lengthy justification to convince sceptical colleagues. Proposals for projects in unusual species are often returned with the suggestion that the applicant use a standard organism instead, because any worthwhile question should be accessible in a well-established model.

Investments in research with a handful of models have returned rich dividends in basic knowledge and medical progress. And many careers, labs and journals are built on the primacy of the fly, mouse and worm[1].

But studying only a few organisms limits science to the answers that those organisms can provide. The extraordinary resolving power of core models comes with the same trade-off as a high-magnification lens: a much reduced field of view. For instance, traditional models for developmental biology — such as the fly — were chosen because their phenotypic traits directly reflect their genotype, with minimal environmental input. These models are poorly suited to questions asked by scientists in emerging fields such as ecological developmental biology — 'eco-devo' — which focuses on external influences on developing phenotypes.

▶

Such limitations have serious consequences. Disparities between mice and humans may help to explain why the millions of dollars spent on basic research have yielded frustratingly few clinical advances[1–4]. Narrowing the research focus too far limits basic understanding, in ways that can lead directly to clinical failures. For example, an experimental treatment for multiple sclerosis that, in inbred mice, improved symptoms of induced disease produced unpredicted — and sometimes adverse — responses in human patients. The inbred mouse model failed to represent the genetic and immunological diversity of human cells, a shortcoming that was obvious in retrospect[2].

It is time to think more critically about how we use models. This means articulating tacit assumptions, such as the adequacy of rodent models to fully represent specific human diseases. It means looking hard at how we select and use our favoured model species, and acknowledging both their strengths and their limitations. And it means mainstream funders and journals welcoming work in non-standard organisms.

> *"To study environmental influences, we need to study species in which such factors matter."*

MODELS OF CONVENIENCE

How did a handful of species become central models? Sometimes it was more about convenience than strategic planning. *Drosophila* rose to prominence in the early 1900s in part because its short generation time was handy for student projects and its four pairs of large chromosomes were ideal for the study of eukaryotic genetics[5]. Yeast, mice, chickens and other domesticated species became lab favourites because they were already familiar and accessible. The existence of lab populations of frogs (*Xenopus laevis*) for use in pregnancy tests led to their recruitment as a model for developmental research.

As model-based science grew, these few species became increasingly dominant, despite the sometimes haphazard way that they had initially been chosen. We have now reached a point where, if researchers cannot tackle a problem using a familiar species, they may not study it at all[1].

Take modern developmental biology. The field has centred on small, rapidly developing organisms with short generation times — most typically, *Drosophila* and *C. elegans*. Much of our current understanding of developmental principles is based on experiments in these species. However, evolutionary selection for rapid development has broad implications. It seems to favour stronger genetic control during development and less plasticity (or flexibility). Compared with related species, development in the models is less responsive to external signals, whether adaptive or disruptive. Because plasticity and the role of the developmental environment are particularly hard to study in key models, these areas receive comparatively little attention[6].

A similar narrowing has occurred in biomedical research. In the case of Parkinson's disease, potential treatments are often assessed by measuring motor function in a lesioned rat. But the rat model does not clearly represent other significant symptoms of Parkinson's that occur in human patients, such as cognitive decline. This may steer some researchers away from these aspects of the disease.

Similar biases rooted in the use of particular models may also contribute to the 'translational disconnect' with regard to neurodegenerative diseases such as Alzheimer's and amyotrophic lateral sclerosis[3,4]. The inability of highly inbred and often genetically modified rodent strains to fully represent the diversity of human patients and symptoms has called the power of such models into question, even within the research communities they serve[1–4,7].

At the same time, the effects of apparently trivial environmental variations, such as the details of mouse handling, are often overlooked[8]. Aggression is the key behavioural phenotype in male mice lacking the enzyme neuronal nitric oxide synthase. This was not observed — and could not be seen — until animals were housed in groups rather than in standard individual cages[9]. Few lab models explicitly account for the environment of organisms, despite increasing recognition that this may affect the outcome and replicability of experiments[7].

In short, if we frame a research model or system too narrowly, leaving out key causal elements such as environmental influences, we cannot hope to construct a complete picture of the mechanisms that underlie crucial variations, for example in development and disease. To study environmental influences, we need to study species in which such factors matter. So the traits that define a successful model must shift as the questions for which we use them evolve.

BEST FIT

Choosing a research model should be more than a matter of convenience or convention. Scientists need to ask more questions — about the goals of a specific experiment, how suitable a given model is to reaching those goals, and what environmental or other external factors might be relevant to how well the model works. For a given question, it is crucial to determine which aspects of human biology are essential (for example, our genetic diversity, unique characteristics of our immune system or particular disease symptoms) and assess how well they are represented in a candidate model (see 'Choosing the right candidate'). Where mismatches appear, we must limit our inferences from animal studies accordingly, and consider when and how to move to research in humans. For some kinds of biomedical

MODEL PROBLEMS

Choosing the right candidate

1 Matching between the model and what it represents

Example: Does studying immunology in highly inbred mouse models shed useful light on the diversity of human immune function and disease[1]?

Key questions

- What do we need to know about a disease to develop treatments?
- What mechanisms link disease origin to symptoms?

Research objectives

- Discover aetiology of symptoms.
- Compare disease initiation and progression between models and humans.
- Assess whether therapeutic targets are well represented in specific models.
- Identify gaps between models and patients that may be significant with respect to basic knowledge and to treatment approaches.

2 Need for additional models

Example: Where there are known obstacles to translating results from mice to humans, how do we develop alternative routes to find new treatments for human diseases[3]?

Key questions

- What aspects of human disease are poorly represented in current models?
- How might the utility of current models be expanded?
- What potential new models are available, or could be developed?

Research objectives

- Develop strategies to assess other aspects of human disease in current models.
- Identify new candidate models for specific questions.
- Develop criteria for selecting new models.

research, it may not matter that the damage or symptoms in the model developed by a different pathway to that which occurs in patients — orthopaedic injuries are one example. But in other areas, such as epidemiology, it matters a great deal.

Recognizing that standard models have limitations does not mean we should give them up. Rather, we should deliberately account for their limitations as part of study design — for example, by analysing the role of a gene in mouse strains with different genetic backgrounds. No single species, no matter how highly engineered, can ever serve as a universal model: every species has unique features that may be assets or faults, depending on the question being asked. For instance, the lack of developmental plasticity in *Drosophila* and of genetic variability in inbred rats limit what these models can tell us about ecological effects on development, but make them powerful tools for studying gene function during development.

We also need to broaden our range of models to include species such as Antarctic icefish, comb jellies, cichlids, dune mice and finches that are naturally endowed by evolution with features relevant to human diseases[10]. Studying the basis of unique adaptive traits in these animals may yield insight into human disorders such as osteoporosis, cataracts and cancer.

Immediately and practically, the US National Center for Advancing Translational Sciences in Bethesda, Maryland, should support the development of new systems for investigating problems that are not tractable in currently favoured models. It should also fund investigations into fundamental questions about model-based research (see 'Choosing the right candidate'). The resulting insights would help scientists to select the best models for advancing basic and applied research, and strengthen the bridges between them. ■

Jessica Bolker *is an associate professor of zoology in the Department of Biological Sciences, University of New Hampshire, Durham 03824, New Hampshire, USA.*
e-mail: jessica.bolker@unh.edu

1. Davis, M. M. *Immunity* **29,** 835–838 (2008).
2. von Herrath, M. G. & Nepom, G. T. *J. Exp. Med.* **202,** 1159–1162 (2005).
3. Geerts, H. *CNS Drugs* **23,** 915–926 (2009).
4. Schnabel, J. *Nature* **454,** 682–685 (2008).
5. Kohler, R. E. *Lords of the Fly: Drosophila Genetics and the Experimental Life* (Univ. Chicago Press, 1994).
6. Bolker, J. A. *BioEssays* **17,** 451–455 (1995).
7. Beckers, J., Wurst, W. & Hrabé de Angelis, M. H. *Nature Rev. Genet.* **10,** 371–380 (2009).
8. Hurst, J. L. & West, R. S. *Nature Methods* **7,** 825–826 (2010).
9. Nelson, R. J. *et al. Nature* **378,** 383–386 (1995).
10. Maher, B. *Nature* **458,** 695–698 (2009).

J. REED/REUTERS/CORBIS

KEEP ABORTION LEGAL
ABORTION IS HOMICIDE www. .org
ABORTION IS HOMICIDE www. ll.org

Pro-choice and pro-life activists clash outside the US Supreme Court in Washington DC.

Politics and fetal diagnostics collide

Without better regulation, non-invasive prenatal genetic tests will be targeted by US anti-abortion lobbyists, argues **Jaime S. King**.

In the United States, pro-life advocacy groups, notably Americans United for Life, based in Washington DC, have been making headway in their mission[1] to limit women's access to abortions "state by state, law by law and person by person". In 2011, 24 US states enacted 92 new provisions restricting abortion — nearly triple the previous record of 34 in 2005 (see 'Clamping down'). One of the strategies of pro-life advocates is to target the reasons for which a woman can have an abortion. Meanwhile, a major development in prenatal care, called non-invasive prenatal genetic testing (NIPT), promises to increase the genetic information available to women early during their pregnancy.

The US Food and Drug Administration (FDA) cannot control how people use information from genetic tests. But by developing a clear regulatory framework for NIPT and improving public understanding of NIPT's benefits and limitations, the agency could help to allay fears that the tests will lead to a drastic increase in selective abortions.

NIPT has the potential to improve women's reproductive autonomy. But if it is not integrated cautiously into prenatal care, the technology could be targeted to support burgeoning strategies to restrict abortion.

In recent years, two blood tests combined with an ultrasound have been the most common method for determining a fetus's risk of having a congenital disease such as Down's syndrome. Results from this type ▶

▶ of test are available only at the beginning of the second trimester. A woman can then choose to schedule an amniocentesis, a more accurate but more invasive test. For this, a clinician inserts a needle into her abdomen to extract a sample of amniotic fluid, which contains the fetal cells needed for genetic testing. The procedure increases the risk of miscarriage by around 1%.

Instead, by analysing fragments of fetal DNA in a pregnant woman's blood, NIPT can reveal potential problems without physical risk. Offered when a fetus is just ten weeks old, NIPT gives a woman much more time to have genetic counselling and confirmatory tests, and to make a reasoned decision about whether to have an abortion while it is still legal for her to do so (in most US states, only before the fetus is 24 weeks old)[2].

NIPT is now used to determine a fetus's blood type, sex and father, and to screen for chromosomal disorders such as Down's syndrome and trisomy 18. The technique is not yet offered commercially for single-gene conditions such as Tay-Sachs disease and cystic fibrosis, but it probably will be soon. Even the use of NIPT to reveal whole fetal genomes may not be far off. In June, researchers sequenced an entire fetal genome from a maternal blood sample[3], and another group did the same a month later[4].

RIGHTS AND REGULATIONS

In the United States, NIPT is emerging just as several states have begun to restrict women's access to abortions sought for certain reasons. In March, a Missouri state representative introduced the Abortion Ban For Sex Selection and Genetic Abnormalities Act of 2012. If this bill were to become law, it would prohibit doctors from carrying out abortions that they knew were being sought because of the fetus's sex or because the fetus had been "diagnosed with either a genetic abnormality or a potential for a genetic abnormality"[5].

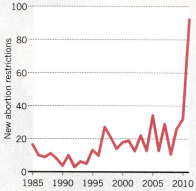
> "The FDA still has not developed a comprehensive regulatory scheme for genetic tests."

Abortions sought because of a fetus's sex are now banned in four states, and bans have been proposed in six others. In May, a bill that would ban all US providers from knowingly performing abortions sought because of the sex or race of the fetus nearly won the two-thirds majority needed, under 'fast-track' rules, to pass the US House of Representatives. Two weeks later, a similar bill, focused on 'sex selection' only, was presented to the US Senate. If it wins majorities there and in the House, it will be sent to the president, whose signature would make it national law.

As the use of NIPT becomes more widespread, pro-life advocates will almost certainly see the technology as a reason to further constrain women's abortion rights. In June, the *National Catholic Register* wrote that pro-lifers view NIPT as "an enhanced 'search and destroy' diagnostic tool" that will drastically increase the number of abortions[6]. Even in Europe, where abortion has historically been a less divisive issue, the technology has prompted anger from various groups. In June, two months before the life-sciences company LifeCodexx, based in Konstanz, Germany, made its PrenaTest for Down's syndrome commercially available, 30 Down's syndrome organizations from 16 countries formally objected to the sale of the test in the European Court of Human Rights.

CLAMPING DOWN

The number of provisions restricting abortions in US states has risen sharply in recent years.



SOURCE: GUTTMACHER INST.

Ideally, no fetus would ever be aborted because of its sex or skin colour. And it is hard to argue that allowing parents to check for hundreds or thousands of traits with one blood test will not facilitate abortions based on societal or individual prejudice. After all, in Asia, there are 160 million fewer girls and women than normal live-birth sex ratios would predict, partly because of the widespread use of ultrasound over the past two decades[7].

But forcing women to have children they do not want will not end prejudice. Instead, it will create a slew of problems. Greater restrictions on abortion may result in more suffering for children[8]. Bills restricting terminations sought for particular reasons will drive a wedge between patients and providers. They will encourage women to withhold information or lie, and they will punish providers serving clients who tell them the truth. Moreover, by dictating which fetuses can legally be aborted, states are entering the dangerous territory of valuing some lives more than others.

US companies that sell NIPT products (such as the California-based firms Sequenom in San Diego, Verinata Health in Redwood City and Ariosa Diagnostics in San Jose) are being cautious. They offer tests only through physicians and for a few conditions, and advertise them as 'screening tests' that may require follow-up procedures. Yet there is too much money to be made from a 'risk-free', relatively inexpensive prenatal genetic test for this restrained approach to last. Worldwide, nearly 50 companies are now developing NIPT products.

HANDLE WITH CARE

The FDA must step up its involvement to ensure that NIPT is integrated into prenatal care carefully — and, especially, to prevent it from being offered directly to consumers, as are other genetic tests.

The FDA still has not developed a comprehensive regulatory scheme for genetic tests, despite repeated calls to do so from government advisory groups (such as the US Secretary's Advisory Committee on Genetics, Health and Society) and non-profit organizations (such as the Genetics and Public Policy Center in Washington DC). This regulatory vacuum is especially problematic in the prenatal context, in which test results can affect parents' decisions to terminate or continue a pregnancy.

The FDA urgently needs to develop a regulatory framework that would allow parents to use prenatal genetic tests under the guidance of a physician and within some general boundaries. As a starting point, the FDA should specify the degree of accuracy and clinical utility required for companies to market a prenatal genetic test. It should also help physicians, pregnant women and the general public to understand the risks, benefits and limitations of such tests — by working with biotechnology companies offering NIPT products, professional societies such as the American Academy of Pediatrics and patient advocacy groups such as the National Down Syndrome Congress.

Abortion has always been a charged issue in the United States. Against this backdrop, NIPT must be handled with care. ■

Jaime S. King *is at the University of California Hastings College of the Law, San Francisco, California 94102, USA.*
e-mail: kingja@uchastings.edu

1. Americans United for Life. *Defending Life 2012* (Americans United for Life, 2012).
2. Greely, H. T. *Nature* **469,** 289–291 (2011).
3. Kitzman, J. O. *et al. Sci. Transl. Med.* **4,** 137ra76 (2012).
4. Fan, H. C. *et al. Nature* **487,** 320–324 (2012).
5. Missouri House of Representatives, 96th General Assembly, 2nd sess. Missouri House Bill no. 1933 (2012).
6. McGovern, C. New Prenatal Testing Could Drastically Increase Abortion Rate. *National Catholic Register* (25 June 2012).
7. Hvistendahl, M. *Unnatural Selection: Choosing Boys over Girls and the Consequences of a World Full of Men* (PublicAffairs, 2011).
8. Sen, B., Wingate, M. S. & Kirby, R. *Soc. Sci. Med.* **75,** 156–164 (2012).

Base sustainable development goals on science

Gisbert Glaser urges the United Nations' working group to do their research.

At the Rio+20 United Nations conference in June 2012, the world's governments agreed to produce a set of sustainable development goals (SDGs). Unlike the Millennium Development Goals (MDGs), which are targeted at poor and emerging nations, the SDGs will have a global reach. They will apply to developed and developing countries alike, and will concern the Earth system as well as people.

This week, governments meeting in New York will discuss SDGs ahead of the launch of the UN working group tasked with defining their scope and path. I call on the representatives of member states to put good scientific data at the heart of the process.

Human pressure on the Earth system may move us beyond safe natural boundaries. As the climate changes, biodiversity is lost and ecosystems decline, we are on course to interlinked environmental, economic and social crises that will make it difficult to provide the growing world population with food, water and energy. Only by setting human development on a sustainable trajectory will we safeguard Earth systems for future generations.

Science was short-changed at Rio+20, where green economy and institutional issues were at the fore. But decisions around complex issues such as water scarcity, ocean health, ecosystems and food security must be evidence-based. As a senior adviser to the International Council for Science (ICSU), headquartered in Paris, I feel it is crucial that the best available research underpins the development of goals, targets and indicators at global, regional and national levels. ICSU, having a special consultative status with the United Nations, has offered to provide scientific input to the working group, drawn from research communities worldwide.

Here I set out the steps that are needed to reach tractable goals that will drive sustainable development. These complement the Future Earth ten-year initiative for global sustainability research, which was launched at Rio+20 by the scientific community.

The SDG idea, put forward by Columbia and Guatemala in 2011, received widespread support at the Rio+20 conference. It builds on the MDG concept of setting voluntary, time-bounded targets. Some people are cynical about these, but I believe that the MDGs, even if not reached, have generated commitments and actions worldwide that would not otherwise have happened. For instance, the MDG of halving between 1990 and 2015 the proportion of people whose income is less than US$1 per day is on target.

GLOBAL VISION

The Rio+20 outcome document[1] proposes that the SDGs must be "action-oriented, concise and easy to communicate, limited in number, aspirational, global in nature and universally applicable to all countries while taking into account different national realities, capacities and levels of development and respecting national policies and priorities".

Meeting all of these requirements will be a challenge for the UN working group. A major difficulty is the interdisciplinary nature of sustainable development. It cuts across economic, environmental and social dimensions in ways that are not well understood. An understanding of climate change, for example, will be necessary to define measures across water, food and energy security. The working group will need to draw on the best available knowledge to analyse these linkages, possible synergies and trade-offs.

The working group's first action must be an extensive information-gathering exercise. This must include all work already undertaken on SDGs, targets and indicators. The group should set up consultations in countries across a range of development levels and seek wide input, from civil society, business, industry and the scientific community.

The long-term strategy of the SDGs must be decided. Should they become the successor to the MDGs after 2015, as some countries have proposed? Or would this direct resources away from unmet MDGs? At Rio+20, all governments agreed that the SDGs should be "integrated into the United Nations development agenda beyond 2015". I agree that is the way forward.

SUSTAINABILITY

Proposed themes

Goals suggested by Colombia, Peru and the United Arab Emirates at Rio+20.

- Food security
- Integrated water management
- Energy for sustainable development
- Sustainable and resilient cities
- Healthy and productive oceans
- Enhanced capacity of natural systems to support human welfare
- Improved efficiency and sustainability in resource use
- Enhanced employment and livelihood security

The goals should be built around cross-disciplinary themes such as food, water and energy security, rather than separate pillars of economy, environment and social development (see 'Proposed themes'). Overarching goals, such as the eradication of poverty, might also be included[2] through a two-tier architecture: a handful of primary SDGs on top of a second layer of thematic ones.

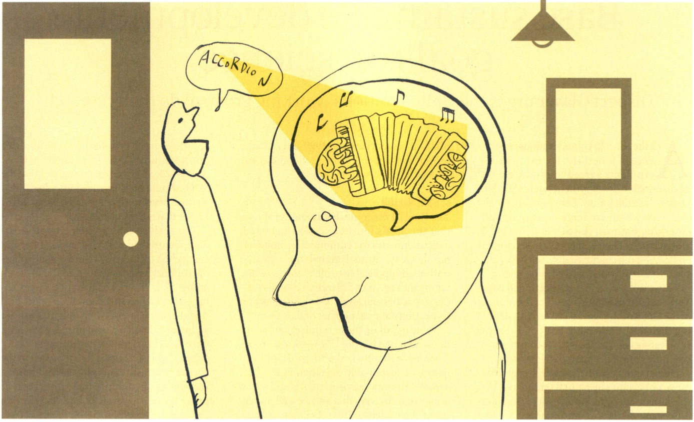
The choice of primary goals is tied to the choice of measures of progress and human well-being. Economic indicators such as gross domestic product and the Human Development Index focus on the short term and fail to reflect the state of the environment and natural resources. Metrics must be based on 'inclusive wealth' — all forms of capital, from natural, social and human to financial and manufactured[3]. As shown by attempts to include ecosystem services and biodiversity within national metrics, this will not be easy.

Measuring progress on the SDGs will require agreed sets of indicators for use at national, regional and international levels and in developed and developing countries. Several have been proposed; none is currently in use everywhere[4,5]. Much more work is needed on science-based indicators.

As consumption accelerates and the world population rises, global sustainability must become a reality. Scientists must help the working group to devise a set of practical targets. Their adoption by the world's governments will be a test of political will. ■

Gisbert Glaser *is a senior science-policy adviser to the International Council for Science in Paris, France.*
e-mail: gisbert.glaser@icsu.org

1. United Nations. *The Future We Want* (UN, 2012); available at http://go.nature.com/ppewxp.
2. Sachs, J. D. *Lancet* **379,** 2206–2211 (2012).
3. UN University International Human Dimensions Programme & UN Environment Programme (eds) *Inclusive Wealth Report 2012* (Cambridge Univ. Press, 2012).
4. UN Department of Economic and Social Affairs *Indicators of Sustainable Development: Guidelines and Methodologies. Third Edition* (UN, 2007).
5. Hak, T., Moldan, B. & Dahl, A. L. (eds) *Sustainability Indicators: A Scientific Assessment* SCOPE Series 67 (Island Press, 2007).



SATOSHI KAMBAYASHI

NEUROSCIENCE

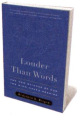
Found in translation

Charles Fernyhough enjoys a bold exploration of how the mind extracts meaning from what we read or hear.

You are doing something quite remarkable. As you read these words, you are taking abstract symbols from the page or screen and extracting meaning from them. They are no longer mere squiggles of ink or pixels — or, in the case of spoken words, patterns of sound. You know what they refer to. Quite how the mind pulls off this nifty trick has troubled philosophers and cognitive scientists for as long as they have been thinking about language.

A prominent view within cognitive science is that linguistic terms are converted into signs or 'tokens' in a 'language of thought', sometimes known as Mentalese. These tokens correspond to the relevant entities in the world. When you read the word 'accordion', for example, a Mentalese token is activated, which allows you to have thoughts about a noisy musical instrument played by squeezing. In his impressive debut, *Louder Than Words*, cognitive scientist Benjamin Bergen tries to persuade us of an alternative view: that we understand language through a process of embodied simulation. Bergen supports this view by reviewing around 200 scientific studies, by his count, from several teams that have been converging on this model during the past couple of decades.

➲ NATURE.COM
For more on linguistics, see:
go.nature.com/i5fbka



Louder Than Words: The New Science of How the Mind Makes Meaning
BY BENJAMIN K. BERGEN
Basic Books: 2012. 312 pp. £18.99, $27.99

According to Bergen's hypothesis, you understand the meaning of a word through the mental recreation of what it would be like to experience the thing being described. So when you hear the word 'accordion', the visual areas of your brain generate an image of an accordion. When you hear the verb 'squeeze', your motor system rehearses the firing that would achieve a squeeze, without going so far as to send the corresponding commands to your muscles.

Much of Bergen's evidence for this account relies on different interference effects, such as the trusty "action–sentence compatibility effect". For example, subjects are asked to read a sentence describing an action (such holding a marble with a closed fist) while simultaneously pressing a button in a way (such as with a flat palm) that is physically distinct from the described action. The mismatch between the described and performed actions slows language processing, suggesting that the comprehension of action-related language shares cognitive and neural resources with the real-life performance of those actions.

An obvious question is how Bergen's system deals with abstracts. Bergen reasons that much of our language about abstract concepts actually relies on concrete metaphors, meaning that both types of language can be underpinned by the same kinds of simulation. For example, studies show that when you read the phrase 'grasp a concept', you comprehend it faster if you have just performed a grasping action with your hand. This theme of complex cognitive abilities being bootstrapped off evolutionarily more ancient systems makes good biological sense. Bergen's view of the language system as a repurposed machine, building on more primitive capacities for perception and action, is one of the attractions of his argument.

But for those familiar with the philosophy and cognitive science of language, there are plenty of unanswered questions in *Louder Than Words*. Explaining how language performs its many functions is notoriously difficult, and the question of how a mental event can ever be 'about' an element of the real world remains a tricky philosophical issue. One problem is that Mentalese isn't the only game in town. You can have an account of meaning based on neural representations of conceptual knowledge that doesn't posit a language of thought. As a result, *Louder Than Words* sometimes has the feel of an assault on a straw man.

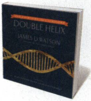
Bergen's treatment of language relies heavily on the 'conduit metaphor', by which one person's thoughts are packaged in words and sentences and unpacked by the listener or reader, rather than shaping and being shaped by a complex patterning of social exchanges. Bergen's theory lacks a developmental perspective: how do babies, for example, get started with word meanings? The model is also unclear on the role of consciousness. We are told that simulation can be unconscious, but elsewhere it is proposed to involve the "creation of mental experiences", presupposing that those experiences are conscious — because, by definition, if you experience something you must be conscious of it.

Bergen ends with a tricky question about the functional significance of mental simulation. Simulation happens when people process language, but does it actually achieve anything? His disarmingly honest conclusion, which dials back on a rather hyperbolic tone at the outset, is that simulation is ultimately neither necessary nor sufficient. He concedes that other forms of language understanding that don't involve embodied simulation are also important.

Bergen admits that his science is still in its infancy, and he sets out his account with enthusiasm, energy and some delightful touches of humour. If you want an engaging, well-informed tour of how cognitive science approaches the problem of meaning, you stand to learn a great deal from this book. ■

Charles Fernyhough *is a reader in psychology at Durham University, UK, and the author of* Pieces of Light, *a book on autobiographical memory.*
e-mail: c.p.fernyhough@durham.ac.uk

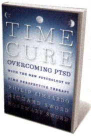
Books in brief



The Annotated and Illustrated Double Helix

James D. Watson, Alexander Gann and Jan Witkowski
Simon & Schuster *368 pp. £19.99, $30 (2012)*

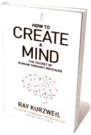
Few tales of modern science thrill as much as the race to discover DNA's double-helical structure. Fifty years after James Watson, Francis Crick and Maurice Wilkins won a Nobel prize for their sprint to the finish, Alexander Gann and Jan Witkowski have crafted a new edition of Watson's behind-the-scenes account, *The Double Helix* (1968). Annotated to clear up abiding mysteries; adorned with lab notes, sketches and photos; and beefed up with extras by Rosalind Franklin and other major players, this is a sampler of rare treats.



The Time Cure: Overcoming PTSD with the New Psychology of Time Perspective Therapy

Philip Zimbardo, Richard Sword and Rosemary Sword Jossey-Bass *336 pp. £17.99 (2012)*

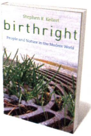
Psychologist Philip Zimbardo has long probed the nature of trauma — notably in his 1971 Stanford prison study — and how orientation towards the past or future affects mental well-being. Now, with therapists Richard and Rosemary Sword, he suggests these findings can guide treatment for people with post-traumatic stress disorder, who suffer harrowing flashbacks. A treatment plan (being tested by the US military), quantitative data and case studies are on offer.



How to Create a Mind: The Secret of Human Thought Revealed

Ray Kurzweil Viking Books *352 pp. $27.95 (2012)*

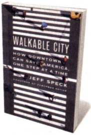
In *The Singularity Is Near* (Viking, 2005), Ray Kurzweil imagined a near future in which medical nanotechnology would allow us to decant copies of our brains into hyper-intelligent machines — and effectively live forever. Now the bestselling futurist and pioneering inventor explores a prime arena for today's big science: reverse engineering the brain. Using the brain's pattern-recognition capacity as a springboard, Kurzweil leaps from the physical brain and the processes of creativity to the debatable idea that, given the correct software, digital entities are effectively conscious.



Birthright: People and Nature in the Modern World

Stephen R. Kellert Yale University Press *264 pp. $32.50 (2012)*

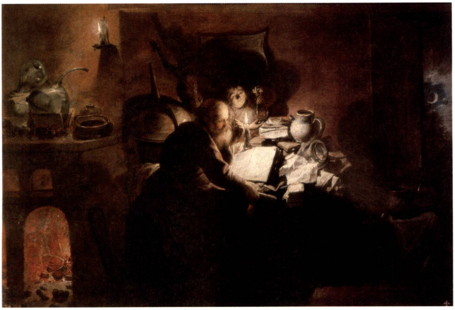
With his sometime-collaborator E. O. Wilson, social ecologist Stephen Kellert has asserted that biophilia — affinity for nature — is central to health, emotional well-being and much more. Here, Kellert challenges our "adversarial" approach to nature with an exploration of eight ways in which we derive meaning from it, from attraction to exploitation. He argues that, even in cities, natural complexity and dynamism are central to children's cognitive development, as they recognize, identify and evaluate rocks, clouds, trees and insects. This is a nuanced analysis punctuated with insightful personal narratives.



Walkable City: How Downtown Can Save America, One Step at a Time

Jeff Speck Farrar, Straus and Giroux *320 pp. $27 (2012)*

Vast intersections, pin-thin pavements, kilometres of concrete: many US city centres were built for cars, not feet. City planner Jeff Speck, working with scores of mayors, concluded that urban liveability demands walkability. He identifies the benefits of well-designed density, such as increased physical fitness, lower fuel use and higher productivity. His 'Ten Steps of Walkability', from well-shaped spaces to curbs on cars, are a blueprint for reclaiming downtown America.



The Alchemist (1640) by David Ryckaert III shows a practitioner at work.

AISA/BRIDGEMAN ART LIBRARY

HISTORY

Realms of gold

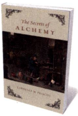
Jennifer Rampling relishes a masterful take on the age-old allure of alchemy.

Around 1680, Robert Boyle, author of *The Sceptical Chymist* (1660), described meeting a stranger who demonstrated an unusual experiment. Tipping some ruby-coloured powder onto the blade of a knife, he cast it into a crucible of molten lead. The lead congealed into "very yellow" metal, which Boyle's tests proved — in his estimation — to be pure gold.

Boyle's account, retold by Lawrence Principe, drives home a problem facing all scholars of alchemy: why, across the ages, have so many intelligent people been convinced by the promise of metallic transmutation? *The Secrets of Alchemy* comes closer than any other single work to explaining the grounds — rational and empirical, as well as religious and wishful — for alchemy's longevity.

Principe's delightful writing style brings to life a depth of learning matched by few in the field. This expertise, coupled with the author's determination to strip his topic of anachronism, sets *The Secrets of Alchemy* apart from the usual introductory tome. After comments on alchemy's lingering popular appeal (think Harry Potter and *Fullmetal Alchemist*), Principe engages with the misconceptions that have long dogged his subject, particularly its association with magic, mysticism and quackery. A key premise of the book is that these are often modern associations. To understand how alchemy 'worked' for its practitioners, we must meet them on their own terms.

Principe traces the theory, practice and context of alchemy from its origins in Egypt in the first few centuries AD to its development and maturity in the medieval Islamic lands and Latin Europe. He then engages with Enlightenment critiques of transmutation, tracing their consequences up to today before returning to alchemy's "Golden Age" in Renaissance Europe.



The Secrets of Alchemy
LAWRENCE M. PRINCIPE
University of Chicago Press: 2012.
296 pp. $25, £16

Some will recognize elements from Principe's earlier work: the argument that 'alchemy' and 'chemistry' overlapped in the early modern world (and so should be referred to simply as 'chymistry'); his concern that Enlightenment polemics and nineteenth-century fads have distorted alchemy's modern reception; and his view that even the alchemists' most outrageous allegories may disguise genuine chemical effects. In sum, he does not believe that alchemists made gold, but does show that they were serious in the attempt.

Like Boyle, Principe recognizes that sceptics will be convinced only by displays of incontrovertible expertise. The book is at its most fascinating when Principe reveals glimpses of his own skill. A chemist as well as a historian, he has recreated a range of alchemical experiments, revealing the practical foundations of seemingly opaque alchemical instructions. The first chapter opens with a recipe from one of the earliest surviving metallurgical treatises, the third-century Leiden Papyrus. The process can be easily replicated, producing a golden patina on a silver ingot. And if Principe's photographic evidence does not convince, an endnote gives instructions on how to do it yourself.

Yet if alchemists engaged in 'real' chemistry, why did they disguise their methods in such baffling ways? In one sequence, Basil Valentine — a fictitious name assigned to a number of alchemical treatises — describes how a king is devoured by a wolf, but resurrected after the creature is cast into a fire. This exotic conceit evolved from an earlier tradition of alchemical secrecy. It also reflects a wider contemporary passion for emblematics — the encoding of meaning in playful images, mottoes and verses.

Such puzzles were designed to be solved. As Principe demonstrates, Valentine's wolfish encounter is an allegory for the technique of purging gold using antimony ore, "the ravenous grey wolf". This process culminates with the volatilization of gold. Difficult enough to achieve in modern labs, this was an astonishing technical feat for early practitioners, who were hampered by impure ingredients, non-standard apparatus and an absence of thermometry. Boyle also cracked Valentine's puzzle, remarking that although difficult, "it eventually succeeded beautifully".

So was Boyle a scientist, alchemist, apologist or interpreter? For that matter, how about Principe? As the book suggests, modern readers can profitably reflect on how they use such distinctions.

For, as Principe concludes, alchemy cannot simply be reduced to chemical procedures. Many practitioners subscribed to a widely held belief in the 'connectedness' of humans, God and nature. In this world view, analogy had demonstrative as well as illustrative power: similarity between small-scale and large-scale phenomena might offer clues to unseen laws of nature. Such correspondences strike us in alchemical writing, because they have disappeared from modern scientific discourse. *The Secrets of Alchemy* reminds that too-selective reading can mask the influence of such views on the past science we now accept as canonical. After Isaac Newton's *Principia*, why not browse his theology — or alchemy? ■

Jennifer Rampling *is a Wellcome Trust Postdoctoral Research Fellow in the Department of History of Science, University of Cambridge, UK.*
e-mail: jmr82@hermes.cam.ac.uk

GEOGRAPHY

Going all the way

Andrew Robinson follows the feet, wheels, ships and space stations that have circled the globe.

In the depths of the cold war, the new nuclear submarine USS *Triton* made the first underwater circumnavigation of the globe in just 84 days. During this covert mission, the crew collected oceanographic and gravitational data, discovered an undersea peak in the mid-Atlantic Ocean and studied the response of human beings to a closed and cramped environment — information deemed useful by the burgeoning US space programme.

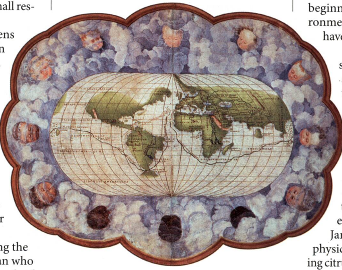
The pioneering 1960 mission deliberately followed most of the route of the first-ever Earth circumnavigation, by Portuguese explorer Ferdinand Magellan and his crew in 1519–22. That expedition limped home after three years, *sans* leader, 86% of its men and four of its five ships. Conversely, the technologically sophisticated *Triton* suffered no fatalities, as its commander (and bestselling novelist) Edward Beach divulged in *Around the World Submerged* (Holt, 1962). But on one 3,200-kilometre detour, it delivered a crewman with a kidney stone to a small rescue party sworn to secrecy.

These two feats, along with dozens more, are expertly filleted in historian Joyce Chaplin's *Round About the Earth*. The book is a first of its kind: a history of circumnavigations covering sea and land, air and space, and almost all forms of transport, from feet and bicycles to Concorde and orbiting space stations. Chaplin makes telling use of details from primary sources. But, as she admits, none of the technologies — whether telegraph, aeroplanes, satellites or the Internet — has, despite grand initial claims, ever "saved the world" on its own.

Choice morsels abound, including the diary of a minor Venetian nobleman who survived Magellan's voyage; the journals of Captain James Cook and his plant-hunting co-traveller Joseph Banks, a future president of the Royal Society; and Charles Darwin's *Beagle* diaries and letters (1832–36). We also get *Following the Equator* (1897) by Mark Twain; the memoirs of solo sailors and fliers such as Joshua Slocum, Francis Chichester and Wiley Post; and stories of the Soviet cosmonauts and American astronauts. Even the celebrated travels of Phileas Fogg in Jules Verne's fictional *Around the World in Eighty Days* rate an extensive discussion.

Chaplin has clearly taken to heart the exchange she quotes from James Boswell's *The Life of Samuel Johnson*. Mulling over the idea of joining Cook's second circumnavigation (in 1772–75), Boswell tests it on the great lexicographer. "One is carried away with the general grand and indistinct notion of a VOYAGE ROUND THE WORLD," says Boswell. "Yes, Sir," Johnson replies, "but a man is to guard himself against taking a thing in general."

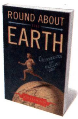
Johnson was prescient, as Chaplin reminds us. For Cook was rewarded less for his geographical discoveries (or his application of the marine chronometer to the longitude problem) than for his essay on the prevention of scurvy, published in the Royal Society's *Philosophical Transactions* in 1776. The British government was deeply impressed that Cook had lost only 2.6% of his crew — and not one to scurvy, which had been the greatest killer of circumnavigating sailors since the time of Magellan.



Cartographer Battista Agnese's 1545 map traces Ferdinand Magellan's expedition.

BRIDGEMAN ART LIBRARY

Chaplin imposes order on her disparate material and draws meaning from it by sticking to the historical chronology of circumnavigation and dividing the half-millennium since Magellan into three major sections: "Fear", "Confidence" and "Doubt". From Magellan's violent death in the Philippines in 1521 to the notorious demise of Cook in Hawaii in 1779, circumnavigators travelled in fear of the dangers. First among these was shipwreck, the result of poor navigational technology and stormy weather, but also starvation and disease at sea, and hostile encounters with indigenous people.



Round About the Earth: Circumnavigation from Magellan to Orbit
JOYCE E. CHAPLIN
Simon & Schuster: 2012. 560 pp. $35

From Cook's death until the Second World War, dramatic improvements in medical treatment, transport and communications technology meant that circumnavigators had much less to fear. Perhaps most importantly, the umbrella of international imperialism endowed these explorers with planet-dominating confidence. HMS *Beagle* carried six guns when it set sail in 1831, whereas by 1872, HMS *Challenger* — a scientific expeditionary vessel exploring the deep oceans — carried a mere two. Moreover, these were intended for signalling rather than self-defence. In the mid-twentieth century, however, the old sense of danger returned, with the development of aeroplanes, rocket-propelled space capsules and the collapse of imperialism. What grew too, argues Chaplin, was doubt: the sense that Earth "is again beginning to bite back, now that the environmental costs of planetary domination have begun to haunt us".

On the whole, the book's tripartite structure works well in the first two sections, but is less convincing in the third, which strangely neglects the issue of global climatic change.

Although science and technology make their presence felt throughout the book, the emphasis is on history, politics, cultures and the personalities of the travellers —the "geodrama" of circumnavigation, as the author calls it. After all, even the scientifically distinguished James Lind (the eighteenth-century physician widely credited with introducing citrus fruits on voyages to protect against scurvy, through the first-ever clinical trial) was convinced that accommodating a sailor's yearning for land was efficacious. Lind reported without a hint of scepticism that scorbutic seamen began to revive when they were taken ashore, stripped and buried up to their necks in the earth. ■

Andrew Robinson *is the author of* The Story of Measurement *and* The Shape of the World: The Mapping and Discovery of the Earth.
e-mail: andrew.robinson33@virgin.net

Correspondence

Handful of papers dominates citation

An 'impact disparity' is emerging in science — only a few papers earn the largest share of citations. This is comparable to the income disparity in the United States, known as the 1% phenomenon, where 1% of the population earns a disproportionate 17.4% of total income (see go.nature.com/yajthu).

The number of citations acquired by a paper is a proxy of its impact. We found that of all papers published in five leading journals in 1990, the most highly cited 1% in each collected around 17% of citations in 2010 (see 'The 1% effect'; data from Thomson Reuters Science Citation Index).

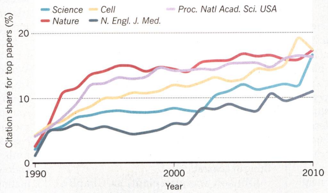
Changes over time in the citation share of the top 1% are evidence of endogenous shifts in underlying processes. These trends are particularly pronounced for citations of older papers. For example, the top 1% of 1990 papers collected only about 5% of citations in 1991.

This shift of attention over time towards the top 1% may reflect the fact that, although the number of research papers has exploded, the time scientists devote to reading them has not. Researchers increasingly rely on crowd sourcing to discover relevant work, a process that favours the leading papers at the expense of the remaining 99%.

Albert-László Barabási, Chaoming Song, Dashun Wang *Northeastern University, Boston, Massachusetts, USA.*
barabasi@gmail.com

THE 1% EFFECT

The top 1% of the most highly cited papers published in 1990 have accrued a disproportionate share of citations over the past 20 years.



SOURCE: THOMSON REUTERS SCIENCE CITATION INDEX

Call to register new species in ZooBank

We wish to clarify a few points in your discussion of the decision by the International Commission on Zoological Nomenclature (ICZN) to allow naming of new species in electronic-only publications (*Nature* **489,** 178; 2012).

The amendment is already in force, retroactive to 1 January 2012 (see, for example, *Bull. Zool. Nomenclature* **69,** 161–169; 2012; available at go.nature.com/atytlr). It remains to be seen which electronic publication will first satisfy the requirements of the amendment. The ICZN official registry, ZooBank, did not support all the requirements until the beginning of September, when the commissioners' votes became official.

New animal species will not need to be registered in ZooBank. It is the electronic works themselves that must be registered to count as published for nomenclatural purposes. In the amendment, registration of new names is a recommendation, not a requirement. We encourage zoologists to comply with this recommendation, which will aid in automated indexing, linking and data extraction.

The amendment allows the ICZN to issue declarations clarifying whether new methods of production, distribution, formatting or archiving can be used to publish works that comply with the requirements of the *International Code of Zoological Nomenclature*. This, coupled with the error-checking capabilities in ZooBank, will enable the code to evolve rapidly and help authors to fulfil the new requirements.

Gary Rosenberg *Drexel University, Philadelphia, Pennsylvania, USA.*
rosenberg@ansp.org
Frank-T. Krell *Denver Museum of Nature & Science, Denver, Colorado, USA.*
Richard Pyle *Bernice P. Bishop Museum, Honolulu, Hawaii, USA.*

Problems enforcing Ecuador ecology law

Ecuador has more species per unit area than any other country, a unique ecology that is now uniquely protected under its constitution. But upholding these highly commendable conservation policies is a challenge.

For example, a landmark legal precedent was set in a lawsuit brought in early 2011 against the local government for damages to the Vilcabamba River caused by a road-construction project. The defendant was ordered to pay for recuperation of the river. One year on, there has still been no substantial remediation (see go.nature.com/6m4aea).

In light of this situation, we are concerned that the imminent strip mining in southern Ecuador of gold and copper ore worth US$200 billion could put a serious strain on the country's legal system and its environmental policies.

Kelly Swing *Tiputini Biodiversity Station, University San Francisco de Quito, Quito, Ecuador.*
kswing@usfq.edu.ec
Luis Sempértegui *Superior Court (retired), Loja Province, Ecuador.*

Open collaboration is key to new drugs

As chair of the board of the Structural Genomics Consortium (SGC), I would like to acknowledge the commitment of the hundreds of scientists from industry and academia who collaborate with the SGC to make freely available synthetic probes that are potentially important to public health. You feature one such molecule, JQ1, now being investigated for blocking unwanted gene expression, in your discussion of epigenetics targets in cancer (*Nature* **488,** 148–150; 2012).

JQ1 resulted from collaboration between SGC researchers and Jay Bradner's group at the Dana-Farber Cancer Institute in Massachusetts, building on the work of scientists at Mitsubishi Tanabe Pharma in Japan and with guidance from scientists at GlaxoSmithKline in the United Kingdom. As you point out, the huge impact of the study is due in large part to the collaborators' willingness to distribute JQ1 without restriction.

Other such examples resulting from open collaborations between industry and academia include inhibitors of the molecules JMJD3 (L. Kruidenier *et al. Nature* **488,** 404–408; 2012), BRD4 (P. Filippakopoulos *et al. Nature* **468,** 1067–1073; 2010) and G9-α-methyltransferase (M. Vedadi *et al. Nature Chem. Biol.* **7,** 566–574; 2011).

Markus Gruetter *University of Zurich, Switzerland.*
gruetter@bioc.uzh.ch

natureOUTLOOK

AUTISM

Produced with support from:
Nancy Lurie Marks Family Foundation, the Simons Foundation, Roche, Autism Speaks, and The Autism Science Foundation

Illuminating a diverse disorder

Roche

www.roche.com

Harnessing emerging science for unmet patient needs

The field of neurodevelopmental disorders is changing fast. Recent spectacular advances in research have revealed that behind clinical heterogeneity there may be common pathophysiological themes, such as synaptic dysfunction in fragile X syndrome and other genetic forms of autism. We now believe autism to be one of the most common severe psychiatric conditions. Autism spectrum disorders are lifelong diseases that severely impact the lives of patients, their caregivers and society, but for which there are very limited therapeutic options. There is a tremendous unmet need for drugs that treat core deficits in social behavior, communication and cognition, which impair the ability of these children to develop, learn and grow into adults capable of having an independent and fulfilling life.

This is a high priority research area at Roche Neuroscience. Scientific breakthroughs in this field have now made it possible to approach drug discovery with a much higher chance of success. This is where our mechanism-based strategy for R&D starts, with a deep understanding of the molecular basis of disease, leading to the assessment of pharmacological targets in translational models, and the development of biomarkers for diagnosis and personalized medicines. In the clinic, our vision is to intervene as early as possible and target both core and associated symptoms to maximize the potential for clinical benefit and disease modification.

Initial results of our research are very encouraging. For example, our scientists recently showed that chronic treatment with an mGlu5 antagonist dramatically reverses the clinical and morphological phenotype in a mouse model of fragile X syndrome.[1] Additionally, in collaboration with academic researchers, we have demonstrated similarities in synaptic dysfunction in genetic and non-genetic models of autism.[2] These findings show that pharmacological therapies can result in disease modifying effects, and confirm that neurodevelopmental disorders may share common tractable pathways for drug discovery.

Together with our partners, we are accelerating R&D in this field. Roche is already making significant efforts to develop therapies for these diseases. Currently, we are investigating three new drugs, in phases 2 and 3, from our internal pipeline in fragile X syndrome, autism and Down syndrome.

The Roche neuroscience leadership team (left to right): Lynda Lowry, Paulo Fontoura, Luca Santarelli, Anirvan Ghosh, Jean-Luc Moreau.

We have also formed key alliances, such as our partnership with Seaside Therapeutics; and our co-leadership with King's College London of one of the largest consortium of academic and industry scientists (the IMI project EU-AIMS) to develop new basic and clinical research tools.

We are committed to developing new therapies for autism and other neurodevelopmental disorders. We are focusing our efforts in identifying new targets and developing drugs, and we know that this will be a long and challenging journey. But with the passion, talent and commitment of our scientists, we are confident that we can develop new groundbreaking therapies to improve the lives of patients, their families and caregivers.

References

1. Michalon, Aubin and *et al.* Chronic Pharmacological mGlu5 Inhibition Corrects Fragile X in Adult Mice. *Neuron.* 2012, Vol. **74,** 1.
2. Scheiffele, Peter and *et al.* Shared Synaptic Pathophysiology in Syndromic and Nonsyndromic Rodent Models of Autism. *Science.* 2012, Vol. **10,** 1126.

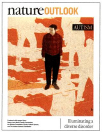
Author

Paulo Fontoura MD PhD
Global Head Neuroscience Translational Medicine, Pharma Research & Early Development, Roche

nature OUTLOOK

AUTISM

1 November 2012 / Vol 491 / Issue No 7422



Cover art: Mark Smith

Editorial
Herb Brody, Apoorva Mandavilli, Michelle Grayson, Tony Scully, Aaron Fagan, Nick Haines

Art & Design
Wes Fernandes, Alisdair Macdonald, Andrea Duffy

Production
Donald McDonald, Yvonne Strong, Kelly Hopkins, Leonora Dawson-Bowling

Sponsorship
David Bagshaw, Yvette Smith, Reya Silao

Marketing
Elena Woodstock, Hannah Phipps

Project Managers
Claudia Deasy, Christian Manco

Art Director
Kelly Buckheit Krause

Chief Magazine Editor
Tim Appenzeller

Editor-in-Chief
Phil Campbell

Imagine you wake up one day and all the familiar signals of human communication are unintelligible. A smile no longer expresses mirth or happiness, a raised voice no longer reflects excitement or anger. You can exchange words with other people, but you're talking through a heavy curtain of uncertainty. You find it difficult or impossible to behave the way people seem to expect. Welcome to the baffling world of autism, a range of disorders that is affecting a growing number of people and continues to perplex scientists searching for causes and cures.

The degree to which autism is on the rise is a matter of some controversy. There's no doubt that the number of children identified as having an autism-related disorder (often described as being on the autism spectrum) has surged in the past decade or so. What's not so clear is whether this represents a true increase in prevalence or just greater awareness of the condition (page S2). Indeed, our understanding of the disorder is evolving — and the shifting definition could deny some children access to the educational and social services that give them a better chance to succeed (S12).

This Outlook is an editorial collaboration between *Nature* and SFARI.org, the news website of the Simons Foundation Autism Research Initiative. SFARI.org operates with editorial independence from the Simons Foundation and did not participate in the editing or commissioning of articles about research funded by the Simons Foundation.

We find that uncertainty still shrouds much of autism. Genetic analysis is beginning to yield candidate genes and some of the underlying physiology of autism (S4). Yet, there are very few treatments available (S14). And little is known about what happens when children with autism grow up (S10).

We acknowledge the financial support of the Nancy Lurie Marks Family Foundation, the Simons Foundation, Roche, Autism Speaks and The Autism Science Foundation. As always, *Nature* retains responsibility for all editorial content.

Herb Brody
Supplements Editor

CONTENTS

S2 EPIDEMIOLOGY
Complex disorder
To define autism we must first figure it out

S4 GENETICS
Searching for answers
Collaborations tackle the complexity

S7 CHILD DEVELOPMENT
The first steps
Study of siblings points to better detection

S10 ADULTHOOD
Life lessons
Research is maturing along with its participants

S12 DIAGNOSIS
Redefining autism
Draft guidelines a matter of contention

S14 TREATMENTS
In the waiting room
Clinicians eager for targeted treatments

S17 PERSPECTIVE
Imaging autism
Nicholas Lange cautions against using brain scans for diagnosis

S18 CULTURE
Diverse diagnostics
Behaviour can only be assessed in context

S20 PERSPECTIVE
Brain scans need a rethink
Evidence for a leading autism theory is flawed, say Ben Deen and Kevin Pelphrey

Nature Outlooks are sponsored supplements that aim to stimulate interest and debate around a subject of interest to the sponsor, while satisfying the editorial values of *Nature* and our readers' expectations. The boundaries of sponsor involvement are clearly delineated in the Nature Outlook Editorial guidelines available at http://www.nature.com/advertising/resources/pdf/outlook_guidelines.pdf

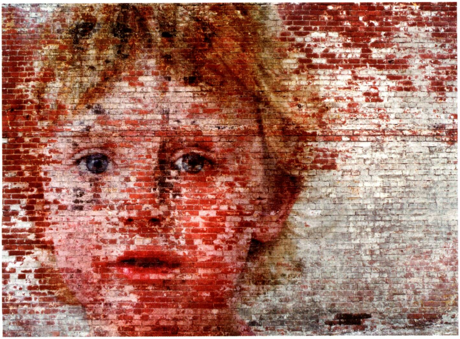
CITING THE OUTLOOK
Cite as a supplement to *Nature*, for example, *Nature* **Vol XXX**, No. XXXX Suppl, Sxx–Sxx (2012).

VISIT THE OUTLOOK ONLINE
The *Nature Outlook Autism* supplement can be found at http://www.nature.com/nature/outlook/autism_2012/
It features all newly commissioned content as well as a selection of relevant previously published material.

All featured articles will be freely available for 6 months.

SUBSCRIPTIONS AND CUSTOMER SERVICES
For UK/Europe (excluding Japan): Nature Publishing Group, Subscriptions, Brunel Road, Basingstoke, Hants, RG21 6XS, UK. Tel: +44 (0) 1256 329242. Subscriptions and customer services for Americas – including Canada, Latin America and the Caribbean: Nature Publishing Group, 75 Varick St, 9th floor, New York, NY 10013-1917, USA. Tel: +1 866 363 7860 (US/Canada) or +1 212 726 9223 (outside US/Canada). Japan/China/Korea:Nature Publishing Group — Asia-Pacific, Chiyoda Building 5-6th Floor, 2-37 Ichigaya Tamachi, Shinjuku-ku, Tokyo, 162-0843, Japan. Tel: +81 3 3267 8751.

CUSTOMER SERVICES
Feedback@nature.com
Copyright © 2012 Nature Publishing Group



CHARLY FRANKLIN/GETTY IMAGES

EPIDEMIOLOGY

Complex disorder

Researchers are digging into the myriad causes of autism to refine its definition and find elusive biological signatures.

BY VIRGINIA HUGHES

Defining 'autism', a single word encompassing a wide spectrum of behaviours, is no easy task.

Each person with autism has a different combination of symptoms, and the combination may change over that person's lifetime. Some people with autism don't speak at all, whereas others are fluent. Some are bothered by sounds, whereas others are musical prodigies. Some are intellectually disabled, whereas others are savants. No wonder then that many researchers have begun referring instead to 'autisms'.

It has become clear over the past few years that the biology underlying autism is similarly multifarious, with hundreds of potential genetic, developmental and environmental causes.

Still, researchers are on the hunt for a biological signature of autism, one they say is the unmistakable core — a distinctive mix of social disinterest and repetitive behaviours — that marks every individual with the disorder.

"That's the big mystery," says David Amaral, a neuroscientist at the MIND Institute at the University of California, Davis. "How could so many different aetiologies, so many different genetic risks, lead to a behavioural syndrome that a clinician could say, 'Well, that person has autism'?"

Since Leo Kanner defined autism in 1943, its symptomatic profile has widened dramatically. It now includes not only social impairments and repetitive behaviours, but many associated ones, such as motor problems, hyperactivity, seizures and distinctive facial features.

Autism's prevalence has also grown. As recently as the mid-1990s it was thought to be rare, occurring in 1 in every 2,500 people. Estimates published in 2012 by the US Centers for Disease Control and Prevention put it at 1 in 88.

Many theories attempt to explain this spike, although no single theory accounts for all of it.

A substantial component of the risk is demonstrably genetic. When one identical twin has autism, for example, there is a roughly 70% chance that the other will also have the disorder.

Many research teams have searched for genes that may be involved. They haven't turned up any prime candidates yet, only dozens, even hundreds, of bit players. "I truthfully believed that sporadic autism, like Kanner had described it, might be caused by a handful of genes with maybe one predominant gene," says Huda Zoghbi, a geneticist at Baylor College of Medicine in Houston, Texas, who in 1999 first identified the gene that causes an autism-related condition called Rett syndrome. "Little did we know."

To complicate matters, many of the genetic blips associated with autism are found in the general population, too. Still, some themes are emerging. For example, many of the genes function at the synapse, the junction between neurons, which has led many pharmaceutical companies to begin early investigations of drugs that target this part of the cell. But the genetic findings are hardly simple. Other implicated genes are related to an array of cell functions, from the immune system, to gastrointestinal repair, and even the production of cellular energy.

Part of the increase must be due to greater awareness. As more people learn about autism, it is likely that more people will spot early signs and seek help. This happens more quickly in some regions than others, depending on cultural and socioeconomic factors. A mix of demographic differences could help explain why, for example, there is a four-fold difference in autism prevalence between Utah, the US state with the highest rate, and Alabama, which has the lowest.

Autism is also diagnosed differently these days. The widening criteria for diagnosis means that many high-functioning people with autism might have gone undiagnosed in the past. Conversely, some of those who have severe cognitive problems might have been diagnosed with mental retardation.

The rapidly rising prevalence has also led scientists to look at environmental causes.

A small part of the increase in autism seems to result from ageing fathers, for example. As men age, their sperm acquires harmful mutations that can be transmitted to the child.

There are any number of other theories, some more plausible than others: the distribution of gut bacteria, autoimmune conditions in the mother, fetal exposure to drugs such as antidepressants, or even living near main roads.

"There have been so many changes in lifestyles," Zoghbi says. "Everything is game."

Despite this seemingly endless list of causes and consequences, researchers are looking for commonalities at every biological level, whether in the genome, a brain circuit or even a style of thinking.

For example, in the 1980s, Uta Frith and colleagues proposed that people with autism lack 'theory of mind', the ability to grasp what others think or believe. Researchers have since localized this ability to a part of the brain called the temporoparietal junction.

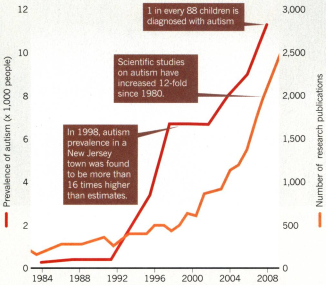
Pairing these kinds of cognitive studies with molecular ones will be crucial to understanding the disorder, says Frith, a developmental psychologist at the Institute of Cognitive Neuroscience at University College London. "You must look at the right level to pull all of the strings together and to talk about this unity of experience that I would call autism." ■

Virginia Hughes *is a freelance science writer based in Brooklyn, New York.*

THE RISE OF AUTISM

Once thought of as a rare disease, autism's prevalence, funding and research has risen dramatically over the past twenty years.

PREVALENCE OF AUTISM AND RELATED RESEARCH

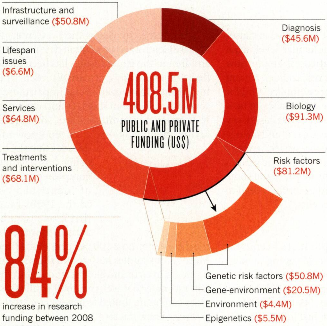


1 in every 88 children is diagnosed with autism

Scientific studies on autism have increased 12-fold since 1980.

In 1998, autism prevalence in a New Jersey town was found to be more than 16 times higher than estimates.

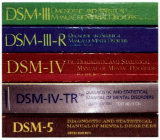
BREAKDOWN OF AUTISM RESEARCH FUNDING IN THE USA, 2010:



84%

increase in research funding between 2008 and 2010

SOURCES: CENTER FOR DISEASE CONTROL AND PREVENTION/ THE INTERAGENCY AUTISM COORDINATING COMMITTEE/BURD, L. ET AL. J. AM. ACAD. CHILD. ADOLESC. PSY. 5, 700–703. (1987)/RITVO, E. R. ET AL. AM J PSYCHIATRY. 146, 1032–1036. (1989)/KIRBY, R. S. J. DEV. BEHAV. PEDIATR. 5, 318–326 (1995)/YEARGIN-ALLSOPP, M. J. AM MED ASS. 289, 49–55 (2003)/ BERTRAND, J. ET AL. PEDIATRICS 108, 1155–11561 (2001).



EVOLVING DEFINITIONS

For just over a century, researchers have grappled with how to define autism — and what causes it.

1911

Swiss psychiatrist Eugen Bleuler coins the term 'autism' in describing self-absorbed adults with schizophrenia.

◆

1943

US psychiatrist Leo Kanner publishes a report of 11 children with autism, defines disorder as "autistic disturbances of affective contact."

◆

1944

Austrian paediatrician Hans Asperger publishes a report of children with profound social problems, lack of empathy and clumsiness.

◆

1967

Bruno Bettelheim's *The Empty Fortress* claims that autism stems from social deprivation, adding fuel to the popular, though incorrect, theory that emotionally cold mothers were the cause.

◆

1980

'Infantile autism' is added to the Diagnostic and Statistical Manual of Mental Disorders (DSM) III. Defined by 6 criteria, including a lack of responsiveness to others, gross language and resistance to change.

◆

1992

Asperger's syndrome becomes a distinct diagnosis when it's included in the tenth edition of the World Health Organization's diagnostic manual.

◆

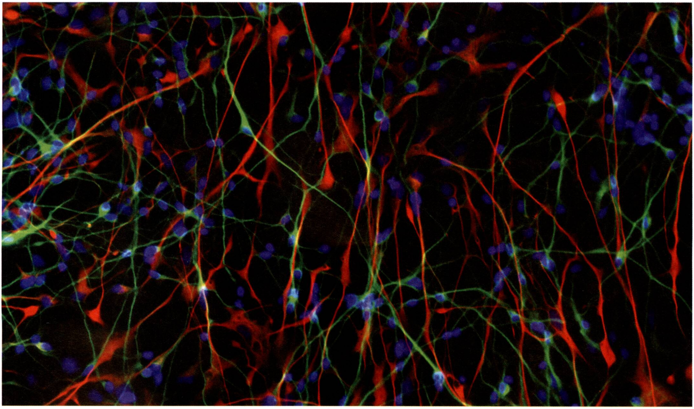
1994

The fourth edition of the DSM greatly expands the autism spectrum, outlining criteria for autistic disorder, Asperger's syndrome and pervasive developmental disorder not otherwise specified (PDD-NOS).

◆

2013

The fifth edition of the DSM is likely to merge the various autism disorders into a single category called autism spectrum disorder.


LUIS DE LA TORRE UBIETA

Researchers are mutating autism-associated genes in brain cells in cell culture to see how the cellular biochemistry is altered.

GENETICS

Searching for answers

Solving the riddle of autism genetics will require looking beyond the growing list of candidate genes to epigenetics and personalized medicine.

BY SARAH C. P. WILLIAMS

At Rutgers University in New Brunswick, New Jersey, blood samples from more than 700 people with autism and their families have been carefully collected and tested over the past five years. The DNA from each of the samples — almost 3,000 in total — has been extracted, studied and shared with researchers around the world. And each person with autism whose genes are in the collection has been put through batteries of tests and examinations to characterize their condition. The Simons Simplex Collection, as the set of data is known, is one of the largest — and most comprehensive — of the handful of autism cohorts around the world. But even so, it hasn't given up many of autism's secrets.

Each new study using the Simons data — or one of the other similar collections — adds to a growing list of gene variations that could relate to the disorder, or perhaps even be the cause in some cases. But on the whole, the expanding catalogue of genes doesn't explain what causes autism at a cellular or molecular level, what characteristics the mutations share, or how to treat the disorder. For researchers, autism is a riddle that is ever more complex. And for patients and their families, it's a story of how science can lead to discovery after discovery but be slow to improve treatments and change lives. Now, however, a new research strategy that integrates protein and epigenetic studies into classic genetic screens might reveal some of the disease's causes and potentially lead to successful, targeted treatments.

THE ROLE OF GENETICS

"From a genetics standpoint, it's been clear for a while now that there isn't going to be an autism gene," says Daniel Geschwind, a neurologist and autism researcher at the University of California, Los Angeles. "But the idea now is that we're moving from just looking under the proverbial street light to integrating many different approaches to solve this problem."

> *"There is just not enough post-mortem tissue that fits our criteria."*

The latest estimates by the US Centers for Disease Control and Prevention state that 1 in 88 children in the United States has been identified with an autism spectrum disorder. Symptoms range from mild social impairment to severe developmental delays and challenges in language, communication, social interaction and behaviour. It's been known for decades that the disorder has a genetic component: studies of identical twins in the 1980s found that if one twin has autism, the other has the disorder almost 100 per cent of the time[1]. More recent studies have found slightly lower numbers in twin and sibling studies, but they still estimate that well over half of all autism cases can be attributed to genetic, rather than environmental, factors[2].

And there's another piece of evidence for inheritance: it's become increasingly clear that older parents — fathers in particular — are at increased risk of conceiving children with autism. A study published in August 2012 found that for every increasing year of a father's age, from 20 to 65 years, the number of mutations passed along to his offspring increased[3]. The research suggests that a man's genetic material accumulates mistakes over his lifetime, as sperm cells are continuously copied. But which of these genetic mutations lead to autism?

SEARCHING THE GENOME

Many of the early studies hunting for autism genes were genome-wide association studies (GWAS). Such studies compare the DNA of people with and without a disorder using thousands or millions of markers scattered throughout their genomes. However, GWAS can only point to large chromosome regions rather than identifying particular genes or gene mutations. The first GWAS on autism, published in 2009, implicated two regions with mild effects: one on chromosome 5, and one on chromosome 20 (ref. 4). Subsequent GWAS on autism have failed to turn up any other parts of the genome with statistical significance. Many researchers were disappointed by these results, which suggested that the common genetic variants that GWAS detects explain only a small fraction of the symptoms of autism. "The common-variant hypothesis was essentially tested," says Geschwind. "And it doesn't mean that common variants don't contribute at all, but they have small effects. So let's look at the rare variants."

But at the University of Southern California in Los Angeles, neurogeneticist Dan Campbell wanted to be sure that all the information possible was culled from the GWAS data before he moved on. "Our approach has been to go back to those genome-wide signals," says Campbell. He was convinced that common variants played some role in autism because of studies that had found broad autistic traits in up to half of all children with autism.

So Campbell's lab started studying, in more depth than ever before, the two regions that had been located. They've already found a gene on chromosome 5 that could play a role. The surprise, and the reason it hadn't been identified earlier, is that it's a non-coding RNA. Instead of being the basis for a protein, the gene is translated into an RNA molecule that has activity in the cell in its own right. In this case, the RNA binds to a different gene — one that makes a protein called moesin, which is known to be involved in brain development. The non-coding RNA turns off the moesin gene, which is critical to early brain development. Without moesin, the neuronal projections (axons and dendrites) are shorter than normal. Brain samples from deceased autism patients, Campbell's team reported in April 2012, have on average 12 times more of the non-coding RNA than usual, suggesting that levels of moesin might have been stunted at some time during development[5]. More work is needed to test whether the altered moesin levels lead to altered brain morphology in people with autism.

Like all the findings on genes implicated in autism, this non-coding RNA doesn't explain all instances of the condition. And even in people who have the common variant and display an increase in these RNA levels, it might not be the full story. "Some cases of autism you can nail down to, 'there's a mutation in this gene'," says Campbell. "Some cases will be due to people inheriting a common genetic variant." And other cases, he said, could be the result of a combination of common genetic variants with an uncommon gene variant or an environmental factor.

With the dropping price of genetic sequencing, many researchers have moved from GWAS to more in-depth sequencing of DNA from people with autism and their unaffected family members in an effort to unearth rarer mutations. In April 2012, three research papers reported the sequencing of exomes (the encoding portions of the genome) from a combined total of more than 1,000 autism patients and hundreds of family members. Together, the studies turned up in excess of 1,000 genes from across the genome associated with autism, far surpassing previous estimates of how many rare genes could contribute to the disorder.

Children born of older fathers are more likely to have autism.

DAVID CRAUSBY/ GETTY IMAGES

The challenge, says geneticist Jay Shendure of the University of Washington in Seattle, is that the rarer a mutation is, the larger the sample size required to prove its association with autism. There are hundreds of thousands of people with autism spectrum disorder in the United States, for example, but only a tiny fraction of that population is part of any well-characterized study like the Simons Simplex Collection. "The cost of sequencing is dropping at such an incredible rate that I do think there will come a time in the next few years when all the good cohorts out there have been sequenced," says Shendure. "But the sample sizes to implicate these genes may need to be much larger, and that means we need more cohorts." The number of people with autism in the United States is large, but it takes time to enrol patients in studies — each person must have in-depth, standardized testing.

Finding the genetic variations is one mission; understanding what they mean is another. Shendure says that new analytical techniques are needed to put the growing list of autism genes into context. By using maps of protein–protein interactions throughout cells, for example, researchers can find networks of proteins affected by many different genetic mutations, suggesting a common mechanism for disease. Or, by relying on new catalogues of gene expression data, they can pin down which cells might be influenced by known mutations. "The more sorts of data we have to intersect our data with, the more we can get out of it," says Shendure.

MOVING BEYOND GENES

The systems biology approach that Shendure proposes is one that Geschwind also endorses. The rare mutations that turn up in exome studies are not necessarily all causative of autism. So using data from the Simons Simplex Collection, Geschwind and colleagues recently profiled the transcriptomes (the set of all RNAs produced) of 244 people with autism and members of their family to get a sense of which genetic mutations led to real changes in gene expression. "Pathway convergence is likely to be quite important," says Geschwind. "You can imagine that many distinct genes could end up leading to a similar circuit-level dysfunction in the brain."

Geschwind's study found that many of these genes with altered expression patterns were clustered in pathways known to play important roles in neuron function, development and structure[6]. This discovery suggests common mechanisms through which seemingly different gene mutations could cause disease. Treatments targeting those pathways might help many patients who have different gene mutations playing into the same networks.

Geschwind's team analysed gene expression patterns in the brains of 19 people with autism and 17 without, using brain tissue collected after each person's death[7]. For each brain, the scientists studied three different regions, all thought to be important in autism. They found that in people with autism, two regions of the brain that normally have distinct patterns of gene expression — the frontal and temporal lobes of the cerebral cortex — instead had almost identical patterns, with the same genes switched on or off. The altered genetics suggest a lack of specialization among some brain cells, which could lead to differences in how the brain processes information.

Geschwind's brain study was small. "There is just not a lot of post-mortem tissue that fits our criteria," he says. "So you are stuck with small cohorts." As tissue collections grow in the future, he hopes to repeat the studies and explore more of the unique brain morphology associated with autism.

At the University of California, Davis, genetics and immunology researcher Janine LaSalle is also looking beyond straightforward genetic

A CATALOGUE OF CANDIDATES

A sample of genetic suspects that may be responsible for causing autism

Gene	Chromosome	Studies that have linked autism to genetic factors	Description
CACNA1H	16	11	Mutations in this calcium channel could change the way neurons function or the way the brain develops.
CNTNAP2	7	20	This gene's protein associates with potassium channels on neurons and may play a role in the differentiation of a nerve cell's axons.
CNTN4	3	9	Encodes a membrane protein that helps axons form in the developing nervous system.
FOXP2	7	16	Encodes a protein that regulates other genes, including *CNTNAP2*. Studies have shown that it is important to neural plasticity.
MECP2	X chromosome	26	The MECP2 protein is implicated in Rett syndrome and known to be involved in turning off other genes in nerve cells.
MET	7	21	Best known as a proto-oncogene, but signalling by *MET* has also been linked to the development of certain parts of the brain.
NRXN1	2	28	Encodes membrane proteins active at the synapse between two neurons.
OTXR	3	20	The OTXR protein helps control the levels in the brain of the neurotransmitter oxytocin.
RAI1	17	7	Mutations in *RAI1* are associated with certain syndromic causes of autism.
SHANK3	22	19	Involved in the development of synapses between neurons. Mice lacking *Shank3* show symptoms.

screens for autism variants, in the hope that she will tie together the hundreds of autism genes being discovered with a finding on another level. Instead of focusing on proteins, LaSalle is looking at epigenetics — changes to gene expression through chemical and structural modifications to DNA, such as the addition of a methyl group to DNA, rather than changes to the code itself.

Epigenetic patterns in cells can be passed from parent to child or influenced by the environment, and can cause major changes to molecular pathways. If epigenetic alterations were discovered in people with autism, they could go so far as to explain the complexity of the disorder, says LaSalle. "Even a subtle hit to one of these pathways through changes in methylation could have the same impact as knocking out an entire gene," she says.

Already, Rett syndrome, an autism spectrum disorder, is providing hints. This neurodevelopmental disorder, which primarily affects females, leads to physical abnormalities as well as repetitive motions and a lack of verbal skills. The syndrome has been linked to a gene, *MECP2,* that controls the epigenetic silencing of other genes in neurons. LaSalle has shown that in mice, mutations in *MECP2* change the way genes are turned on and off in response to a chemical found in flame retardants. Her results don't prove that these chemicals are a cause of autism, but they suggest that *MECP2*'s epigenetic mechanisms could link genetic and environmental impacts in autism[8]. Researchers in the emerging field of environmental epigenetics have already found that many environmental toxins and pollutants reduce overall methylation in cells, she says.

But epigenetics is a young field and new techniques are needed to be able to fully understand how methylation is controlled and, in particular, how it might have unique roles in neurons. There's evidence that in neurons, methylation is influenced not only by environmental and inherited factors, but also by neuronal activity itself, setting up a feedback loop. Moreover, neurons have been found to show different epigenetic patterns to the rest of the body[9]. Until recently, it was thought that most methylation was concentrated in the promoters of genes, so most commercial tests for methylation are biased toward promoters. As the technology for studying these patterns improves, LaSalle says, additional epigenetic associations with autism will emerge. For example, epigenetic modifications could influence the same genes and pathways that genetic mutations alter. "The real key is overlaying the epigenetic data with data from genome-wide association studies and having people work together," says LaSalle.

> ***"I see cancer as being about a decade ahead of neuroscience."***

OPTIMISTIC HUNTERS

So far, no single gene has been discovered that accounts for more than a few per cent of autism cases. Still, the discovery over the past decade of more than 100 autism-related genes is progress of a sort. Indeed, according to Geschwind, known diseases and mutations now collectively explain almost one-fifth of autism cases. "If you asked me 10 years ago if I thought we'd be on the threshold of understanding some genetic risk in 20–25% of autism, I would have thought that would be very, very optimistic," he says. "But the field has moved rapidly."

Most recently, a potential treatment for a rare form of autism that involves a mutation in the *BCKDK* gene was reported[10]. Experiments with *BCKDK*-deficient mice showed improvement when administered diets rich with branched-chain amino acids (BCAA). *BCKDK* mutations diminish BCAA levels, and in human patients the regimen helped restore plasma BCAA levels. These are promising developments that could lead to treatment by sequencing exomes in order to link specific genes to their physiological effects.

Eventually, Geschwind says, treatments for autism may become personalized. Genetic tests could determine which treatments would work in a certain patient's case. But first, studies on broader protein networks and the underlying causes of autism are likely to lead to more general treatments that work for at least some patients, although as these treatments will target broad cellular pathways, they might bring unwanted side effects. Geschwind cites examples from other clinical sciences. "I see cancer as being about a decade ahead of neuroscience," he says. "In cancer they've been effective in some cases but have also seen off-target effects. In many ways that field is now moving towards more targeted drugs."

The large databases of autism gene candidates that are now available make the quest to explain autism more complicated than researchers had hoped. But the complexity of the condition is stimulating the expansion of approaches taken and enticing scientists to look beyond straightforward genetic explanations for autism. "We've figured out that explaining autism is not simple," says Geschwind. "But I have a pretty optimistic view. We're going to continue to make progress — and a lot of it is because of great collaboration in the field and an influx of new people tackling autism." ■

Sarah C. P. Williams *is a freelance science writer based in Kailua, Hawaii.*

1. Ritvo, E. R. *et al. Am. J. Psychiatry* **142,** 74–77 (1985).
2. Ronald, A. *et al. J. Am. Acad. Child Adol. Psy.* **45,** 691–699 (2006).
3. Kong, A. *et al. Nature* **488,** 471–475 (2012).
4. Weiss, L. A. *et al. Nature* **461,** 802–808 (2009).
5. Kerin, T. *et al. Sci. Transl. Med.* **4,** 128ra40 (2012).
6. Luo, R. *et al. Am. J. Hum. Genet.* **91,** 38–55 (2012).
7. Voineagu, I. *et al. Nature* **474,** 380–384 (2011).
8. Woods, R. *et al. Hum. Mol. Genet.* **21,** 2399–2411 (2012).
9. Iwamoto, K. *et al. Genome Res.* **21,** 688–696 (2011).
10. Novarino, G. *et al. Science* doi: 10.1126/science.1224631.

STACEY WIESNER

A child at low risk for autism wears an electrode cap, which captures the faint electrical signals generated by neurons, as part of a study on autism development.

CHILD DEVELOPMENT

The first steps

Because infants born into families with autism are more likely to develop the condition, studying them might lead to ways to diagnose people in the general population earlier.

BY KATHERINE BOURZAC

At six months of age, Case 6 was a happy baby. She locked eyes with people and smiled back, reacted when her name was called, and enjoyed playing 'peek-a-boo'. By twelve months, she babbled and knew three words. Six months later, she hadn't learned any new words. She didn't point at things to ask for them, as normally developing babies do at this age. She fell frequently and began hitting herself on the head. At 36 months of age, she was diagnosed with autism.

Many parents whose children have been diagnosed with autism can tell a similar story. The individual details are different, but the global progress of the disorder is the same, says Lonnie Zwaigenbaum, now co-director of the Autism Research Centre at Glenrose Rehabilitation Hospital in Edmonton, Canada, one of the leaders of the study that included Case 6. The girl labelled 'Case 6' was part of a study that followed, from infancy to toddlerhood, the younger siblings of children diagnosed with autism. Baby sibs, as researchers affectionately call them, are at a much higher risk of developing the disorder than the general population. One in five baby sibs will be diagnosed with autism[1]; the prevalence in the general population is 1 in 88.

Because the population of baby sibs is almost 20 times richer in children who will go on to develop autism, this cohort is ideal for studying the origins of the disorder. In baby sib studies, researchers record the behaviour, cognitive ability and brain function of baby sibs throughout development, as well as infants the same age who do not have any siblings with autism. By comparing the developmental trajectory of the two groups, they hope to gain a better understanding of how autism unfolds. But more importantly, researchers are on the hunt for markers that might enable them to identify infants at high risk of autism in the general population, while developing earlier interventions.

The story of babies like Case 6 does not have to be set in stone at birth. "There is no reason to accept the disabling consequences of autism," says Mayada Elsabbagh, a psychiatrist at McGill University in Montreal, Canada, who studies the brain function in autism. "If we start intervention earlier, when the brain is more malleable, we can prevent these consequences."

EARLY SIGNS

Autism can be diagnosed at three years of age. However, owing to disparities in access to healthcare, the median age at diagnosis in the United States is more than five years. By the time most children with autism come into the clinic, they already have major problems with social engagement and language, as well as other difficulties, such as repetitive behaviours.

Early intervention is important because autism can create a vicious cycle. Without the ability to draw people in and be drawn to them, a child will lose learning opportunities that shape the rapidly developing brain. But cause and effect are difficult to tease apart in autism, and problems in brain function are likely to precede social and language difficulties. As a

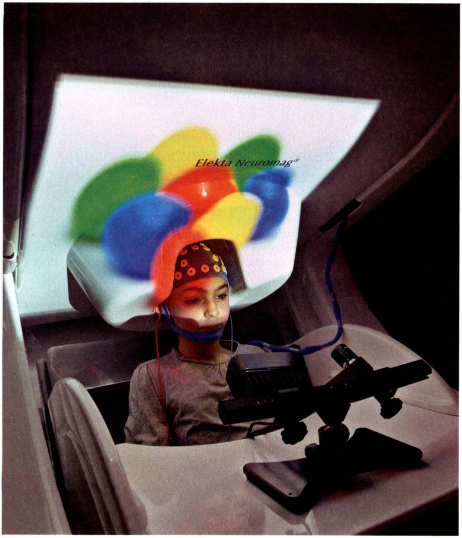
➲ NATURE.COM
Unlocking the brain's neural circuitry to treat brain disorders:
go.nature.com/2npoxd

result of this muddle, "by the time you see the end state, it's hard to separate the core symptoms from their effects," says Mark Johnson, who directs the Centre for Brain and Cognitive Development at Birkbeck, University of London. Prospective baby sib studies might fill in the gap.

In the mid-1990s, researchers started the first prospective studies that followed both baby sibs and children with no siblings with autism. The earliest of these studies focused on behavioural markers. In 2005, having followed a group of infants (including Case 6) from six months of age, Zwaigenbaum co-authored one of the most influential baby sib studies to date[2]. Based on this investigation, Zwaigenbaum says, while there is a clear global progression of behaviours in baby sibs who go on to develop the disorder, when each infant is watched individually, there are exceptions.

At six months, most of the babies who go on to develop autism are in line with their peers. Like Case 6, they smile when their parents smile, and when strangers smile. Most will perk up when someone says their name. However, while most babies around this age develop the strength and coordination to hold up their heads and sit independently, baby sibs who go on to develop autism often do not. "They may be floppy — some can't hold their head up," says Zwaigenbaum. These early motor problems are one of the best-established early indicators of the disorder, but why they occur, or whether they might cause subsequent problems, is unknown.

After six months, the development of baby sibs who go on to develop autism seems to plateau, whereas their peers become more social, more talkative, and more creative. By 12 months, the Canadian researchers found, many children who will eventually be diagnosed with autism have started to become withdrawn. They don't cue into what other people are doing, but may be intensely interested in non-social stimuli. For example, they may stare at toys but not play with them in ways other children do. "They don't smile as much, they don't light up to peek-a-boo, and they show intense visual engagement with toys," says Zwaigenbaum. "But it's variable — you might find they are reserved but not impaired, and they may score normally on behavioural assessments."



Eye trackers and magnetoencephalography scanners detect neural activity and are helping to quantify the early cognitive signatures of autism.

JAMES KING-HOLMES/SCIENCE PHOTO LIBRARY

Children who will later be diagnosed with autism start to show more stable, recognizable signs after 18 months. "They make poor eye contact, they don't play [normally], and they begin repetitive behaviour," says Zwaigenbaum. At this age most normally developing children begin pretending with toys, but a child who goes on to develop autism is less likely to. Instead of playing imaginatively, the child may, for example, put the wheels of a toy car up to his or her face and stare intently as they spin.

Although these behavioural and motor signs are clearly correlated with an autism outcome, says Zwaigenbaum, neither his group of children nor any others have demonstrated that they can be used to predict the disorder. Yet even though the results are confounding, they also offer hope. Baby sibs who do not go on to develop the disorder often score as atypical at some point, compared with infants who have no siblings with autism. This suggests that for some baby sibs — who are presumably at higher genetic risk for the disease — there may be protective factors at work that researchers could identify and encourage.

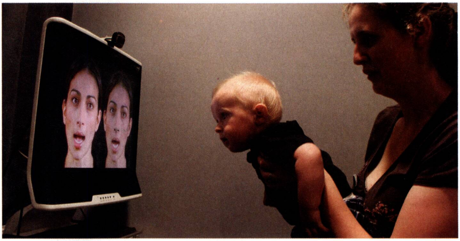
QUANTIFYING RISK

A new crop of baby sib studies goes further and deeper. Ami Klin, chief of the Marcus Autism Center in Atlanta, Georgia, is trying to fill in autism's developmental map with dense, quantitative data on behaviour. His approach is to measure early and often. He believes this is the only way to get at the richness of normal development in the first year of life, and how it goes awry in autism. In his group's studies, the children are assessed monthly for their first 6 months, then every 3 months until they reach 24 months of age. Measuring once every 6 months, he says, is like "sending a spaceship to another planet, taking a picture from far away, and saying, 'no water'," — without ever sending a rover for on-the-ground exploration.

Klin's lab is developing technologies to quantify infant attention. One of these is an eye-tracking system that very densely samples where infants look, as often as 60 times a second, with software to analyse the data. Klin developed these tools and tested them in older children with autism. For example, Klin's group reported[3] an eye-tracking test that compared how attentive toddlers with and without autism were to social and non-social motion by tracking them while they watched simple animations. Toddlers without autism preferred to look at human motion, whereas those with autism showed no preference.

Klin has now done similar vision-tracking experiments in baby sibs, although the results haven't been published yet. He is waiting for the children to reach 36 months, when the autism diagnosis is considered stable, so he can compare baby sibs who did and did not develop the disorder with the control group.

HELEN MARIS



You looking at me? Whether infants repond to a direct gaze is one of the tests for autism being evaluated.

He's also working with speech scientist Gordon Ramsay of the Yale Child Study Center in New Haven, Connecticut, to make monthly recordings of the babies' vocal environment, to watch how they progress from babbling to language, and how they learn to modulate their voices to convey emotion.

THE CHICKEN AND THE EEG

Behavioural checklists and quantitative measures of cognition such as Zwaigenbaum's and Klin's may eventually lead to a test that paediatricians can use to identify high-risk babies in the general population. To complete the developmental map of the disease, other researchers are undertaking studies of brain function. But there's still a chicken-and-egg problem in autism: do problems in brain function precede behavioural problems, or vice versa? And to what extent do these feed off one another?

"Changes in brain development precede changes in behaviour," states Charles Nelson, director of research for the Developmental Medicine Center at Boston Children's Hospital in Massachusetts. Nelson's is one of a handful of groups studying brain function in baby sibs. Since most behavioural signs of autism emerge after the first year, Nelson and others believe that the key will be to find the origins of autism in the brain during the critical first 12 months of life. Early results from a handful of these studies have come out this year, and many more will follow.

One of the most powerful tools available to developmental neuroscientists such as Nelson is electroencephalography, or EEG. Researchers fit babies with electrode caps that capture faint electrical signals generated by neurons firing on the surface of the brain. In 2012, Nelson's group showed that 6-month-old baby sibs, when resting, showed lower levels of all types of EEG activity compared with low-risk babies[4]. His group is still processing data from older ages.

In the spring of 2012, researchers led by Johnson and Elsabbagh showed that brain function between 6–10 months of age could be used to predict which baby sibs will be diagnosed with autism at 36 months of age[5]. While monitoring them with EEG, the researchers showed the babies two images: one of a woman looking at the viewer, and one with her eyes averted. Most of those who did not develop autism showed a different EEG response depending whether the woman's gaze shifted to or away from them. Babies who were later diagnosed with autism did not appear to differentiate.

In these brain studies, as in the behavioural studies, the clearest differences were not between infants who developed autism and those who didn't, but between high-risk infants — the baby sibs —and low-risk infants. "Most babies at risk are processing stimuli differently, but many compensate for that" and do not develop the disorder, says Johnson. Perhaps they compensate because the infant brain is plastic, or perhaps they have a particularly supportive environment. Whatever it is, it's giving researchers hope that there might be protective effects — and that early interventions could eventually alleviate autism.

It may be a decade or more before behavioural and brain-function studies in baby sibs yield a reliable checklist that can be used to predict the risk of autism in the general population. For one thing, researchers can't crunch the numbers until three years into the study, when the children reach an age at which a stable diagnosis of autism can be determined. And then, once the data from these prospective baby sib studies are sufficiently robust, researchers will then have to use them to test babies in the general population and follow them until 36 months as well.

In the meantime, some researchers are testing infant interventions in baby sibs even without having a reliable prediction of autism. Because only one in five of these children will receive a diagnosis, any intervention has to be supportive and positive for both the infant and the parents, regardless of whether the child develops autism. Programmes in testing emphasize fostering the relationship between parent and baby — something that's good for any family, and that may have particular benefits for children who will go on to develop autism, perhaps lessening their eventual symptoms, says Jean Kelly, co-director of the Center on Infant Mental Health and Development at the University of Washington in Seattle.

Whether improving parent–baby relationships works to lower the risk of autism is the subject of baby sibs intervention studies at the University of Washington and elsewhere. Kelly's group takes videos of parent–child interactions, and coaches parents — many of whom are stressed because they already have a child with autism. Kelly holds these sessions weekly for 3 months with baby sibs 8–15 months of age. Kelly's group is trying to help teach, "How do I reach *this* child?" And, she says, "we are promoting trust, security and attachment on the side of the child."

Rebecca Landa, director of the Center for Autism and Related Disorders at the Kennedy Krieger Institute in Baltimore, Maryland, has seen the power of interventions in high-risk toddlers from two years of age and is now trying to use these measures in babies. Some of the interventions her group is studying involve bringing together 12-month-old baby sibs who have shown troubling behavioural signs. The group then performs social tasks that take advantage of the peculiarities of repetitive autistic behaviours to encourage them to imitate one another — something these children normally won't do. These children tend to move objects back and forth and stare at them, so if the researchers give them all the same object, they end up doing the same motion, and then they notice one another. In one exercise, the 1-year-olds wipe the table in front of them with wet wipes, and the researchers then get them to sing "This is the way we wash the table" and carry the wipe to the bin.

After this exercise, says Landa, "the parents are in disbelief that the children are doing an activity they normally do at home, that they're learning the words 'wash table', that they are socially engaged," she says. The intervention studies are still in progress, but Landa is optimistic. "We want to stop these problems in their tracks," she says. If these methods are shown to have benefits for baby sibs, Landa and Kelly both hope to develop a standardized program that could be widely implemented by paediatricians and child minders.

"For some of these babies, autism doesn't have to be their destiny," says Johnson. Genes influence the path of development, but the course of development is not inevitable. ■

Katherine Bourzac *is a freelance science writer based in San Francisco, California.*

1. Ozonoff, S. *et al. Pediatrics* **128,** 488–495 (2011).
2. Zwaigenbaum, L. *et al. Int. J. Dev. Neurosci.* **23,** 143–152 (2005).
3. Klin, A. *et al. Nature* **459,** 257–261 (2009).
4. Tierney, A. L. *et al. PLoS ONE* **7** (6), e39127 (2012).
5. Elsabbagh, M. *et al. Curr. Biol.* **22,** 338–342 (2012).

NLM
FAMILY FOUNDATION

Grounding Autism in the Basic Biomedical Sciences

The Nancy Lurie Marks Family Foundation (NLMFF) seeks to enhance the lives of people with autism by funding basic biomedical research, providing resources for advocacy and policy groups, and making available social services for families. We believe that access to improved medical diagnosis and treatment and to better means of communication is the surest path for autistic individuals to find increased enjoyment in social and family life, obtain meaningful employment, and discover ways of independent living and creative expression.

Achieving these goals means that autism must be approached from a fundamental scientific perspective. It is only when autism comes to be understood and defined in terms of off-kilter physiological systems and circuits, and not just their surface manifestations, such as self-injurious behaviors, seizures, tics, aggression, and frustratingly futile attempts at speech, that stigmatization and isolation will end. We believe that a proper description of autism will set the stage for powerful new technologies to be brought to bear on the difficult problems these children and adults face. Medical history has shown repeatedly that basic knowledge leads to practical invention, often in completely unexpected ways, but for this to happen in the case of autism the challenge needs to be expressed in terms of targetable physiological or molecular systems and processes, not just behaviors, failures to communicate, or socially awkward habits.

The principal goal of the scientific program at the NLMFF is to contribute to this deeper understanding of the biological basis of autism. The astonishing discoveries of the past few years in neuroscience allowing, for example, the electrical properties of individual neurons to be manipulated in living animals, enabling signaling properties of genotype-matched cells to be grown in culture, or revealing the subtle correlations in frequency coupling across brain regions as new tasks are learned, place demands on our scientific imaginations to probe the nature of autism more deeply than what appears on the surface in the form of behaviors.

Science is a conversation, across generations and disciplines, and for too long autism research has been isolated. To help bring autism into the conversation of modern neuroscience, NLMFF conducts small workshops to which leading scientists, not necessarily engaged in autism research, are invited to share their insights with workers in the field. These meetings often inform our funding decisions and serve as an impetus for young scientists to enter into autism research.

Much of our funding at NLMFF is focused on genetics, synaptic chemistry, systems biology and the physiology of movement, all with the aim of exploiting the new insights of neuroscience to discover how behavior is organized, coordinated, and controlled in the human body, and how language is learned and speech produced. The scientific study of communication modalities, hand, eye, and ear, with the goal of producing new and useful devices and strategies for social and intellectual engagement, is essential for developing support systems that could enhance the quality of life for many children and adults with autism by alleviating a principal source of distress. Autism is the 'Rosetta Stone of human neurobiology', presenting rich opportunities to study the neurobiology of speech, brain rhythmicity, and emotion, and explore the nature and potential of human thought and creativity as it evolves on a different, but no less endowed, developmental trajectory.

LURIE CENTER FOR AUTISM

Nancy Lurie Marks and the Nancy Lurie Marks Family Foundation recently made a gift to the Massachusetts General Hospital (MGH) to establish the Lurie Center for Autism, an integrated and multidisciplinary clinical, research, training and advocacy program dedicated to treating individuals with autism spectrum disorders and other developmental disorders across the lifespan. In addition to clinical care, cutting-edge translational research, advocacy and public policy analysis, the Center is committed to providing training and clinical experience for a new generation of medical doctors to meet the comprehensive needs of those with autism.

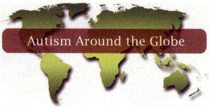
Persons with autism, like the rest of us, face new and different medical challenges as they age. Little is known about what health concerns might be unique to this population. Medicine has always advanced from the hunches and conjectures of physicians who see hundreds of patients, often for periods spanning many decades, and large clinics devoted to serving this population are needed.

To address these and many other questions, the Lurie Center will maintain comprehensive longitudinal biomedical databases, integrating clinical and research observations, enabling qualified researchers to assemble research cohorts with informed consent. Inherent in these databases may be patterns and subgroupings that will spark physician-researchers,

LURIE CENTER FOR AUTISM

MASSACHUSETTS
GENERAL HOSPITAL

MassGeneral Hospital
for Children™



collaborating with bioinformatics specialists, to form new hypotheses and generate new ideas for treatments. Through referrals to specialists at the major clinical departments at MGH, as well as collaborations with basic scientists at the Martinos Center for Biomedical Imaging and the Center for Human Genetics Research, the Lurie Center is in position to ground the study of autism in one of the most advanced biomedical treatment centers in the world.

In addition to its core services, the Center offers speech, occupational, and physical therapies to autistic individuals. The clinical program, specializing in diagnosis and referral, includes experts in pediatrics, adult internal medicine, gastrointestinal evaluation, augmentative communication, nutrition, audiology, and optometry. Large hospital systems can sometimes present daunting challenges for persons with autism just to get to the right place, compose oneself, and then try to communicate what's not feeling right. Families are assigned a dedicated 'Autism Resource Specialist', often a medical social worker, a person who will come to know them well, to help with the planning and coordination of their visits with clinicians, therapists, researchers and school counselors.

NANCY LURIE MARKS CLINICAL AND RESEARCH FELLOWSHIP PROGRAM IN AUTISM

If the field of autism is to be transformed by the discoveries of biomedical science in pediatrics, neurology, and psychiatry, in addition to the basic sciences of genetics, physiology, and structural biology, training in autism must become integrated into the curricula of medical schools. Physicians specializing in autism are in precious short supply. To help meet this need, and increase awareness of the difficulties autistic individuals may face in the clinic, The Nancy Lurie Marks Clinical and Research Fellowship Program in Autism has been established to support Harvard Medical School faculty and students interested in autism and related neurological disorders. Participants are selected from all levels of the training program: medical students staying on for a fifth year, advanced postdoctoral scientists seeking clinical collaborations, and doctors who have completed their residency and are undertaking specialty fellowship training.

LURIE INSTITUTE FOR DISABILITY POLICY

We are also keenly interested in addressing progressive public policy and planning initiatives relevant to adults with autism including employment, residential planning, family and community supports, and the transition to more independent living. Through a grant to Brandeis University, the NLM Family Foundation helped establish the Lurie Institute for Disability Policy and endow a professorship at the Heller School for Social Policy and Management. The Lurie Institute for Disability Policy conducts research on disability policy in the United States with a special emphasis on autism, focusing on the lifespan of persons with disabilities and their families. The Institute explores how human services and public benefits are provided to persons with developmental disabilities and their families and also how persons with disabilities can become self-advocates. By offering courses on disability policy, the Institute fosters and develops careers in this field. The Institute also shapes public policy on issues related to autism and other disabilities as these issues affect children, the aging population, special education, health care, and transitions across the lifespan for persons with disabilities and their family caregivers.

AUTISM AROUND THE GLOBE

Autism around the Globe (www.autismaroundtheglobe.org) is a global public awareness project of the NLM Family Foundation. The goal of this project is to provide a platform for sharing stories from contributors around the world whose lives have in some way been affected by autism. These compelling narratives, often translated into multiple languages, provide insight into the cultural factors that affect those living with autism in different parts of the world. There are many surprising and original observations on how autism is perceived in highly developed and populous areas as well as in remote regions of the world. Detailed descriptions of referral and diagnostic procedures in places where medical care is so different from that typically found in advanced countries vividly illustrates the universality and devastating impact of autism on daily life and the long-term prospects for adults living with this condition. Many of the authors of these articles are medical researchers or directors of caregiving agencies, who are passionately interested in joining the world-wide research effort to understand how autism arises and what can be done about it.

AN INVITATION TO AUTISM RESEARCH

The pace of discovery in autism research now matches that in many other areas of clinical neuroscience, including schizophrenia, addiction, Alzheimer's disease, and depression. In some cases, such as the convergence onto common synaptic mechanisms involved in learning and habit formation, the large number of implicated genes has, remarkably, led to simplifying and testable hypotheses at the systems level. The Nancy Lurie Marks Family Foundation encourages scientists, young and old, basic and applied, to take up the challenge of understanding autism and, in the process, discover how molecules, cells, and circuits give rise to learning, memory, and human communication.

Author

Clarence E. Schutt, Ph.D.

Director & Chief Scientific Officer,
NLM Family Foundation,
60 William Street, Suite 110,
Wellesley, Massachusetts 02481

Professor, Department of Chemistry,
Princeton University,
Washington Road,
Princeton, New Jersey 08544

JOHN BIRDSALL/PHOTOFUSION

ADULTHOOD

Life lessons

We know little about autism past adolescence, but a well-studied generation of children with autism will change that.

BY LINDSAY BORTHWICK

The first child ever diagnosed with autism, by child psychiatrist Leo Kanner in 1943, was five-year-old Donald Gray Triplett. Described by his father as living "within himself" with "no apparent affection" and a "mania for spinning blocks", Donald would become the archetype for a developmental disorder that is now one of the most common in the United States — affecting an estimated 1 in 88 children, according to the US Centers for Disease Control and Prevention.

These days, though, Donald lives much like any other American retiree. A 2010 profile in *The Atlantic* magazine portrays him as a happy, independent, Cadillac-driving septuagenarian — albeit one with a few quirks, including a prodigious gift for numbers — who plays golf and travels the world.

The contrast between Donald as a young boy and an old man is striking. The question is, how did he get there?

We don't know, says Joseph Piven, a psychiatrist at the University of North Carolina at Chapel Hill, who in 2010 spearheaded a working group on adults older than 50 years of age with the disorder. Older adults with autism, he says, "are virtually unstudied".

It is even unclear whether the disorder's core symptoms — social and communication deficits, repetitive behaviours and restricted interests — persist during adulthood. Nor do we know much about the prevalence of autism among adults or about their mental state, brain activity or general health.

"Early in my career, if you said this is a lifelong condition, people would get upset. They didn't want to hear that," says Piven. But that perception of autism as a childhood disorder has begun to change, largely due to shifting demographics. Autism prevalence has risen sharply since the 1990s, so much so that neuroscientist and psychiatrist Thomas Insel has likened it to a "huge wave moving through the system". Insel is director of the US National Institute of Mental Health in Rockville, Maryland, and chairs the Interagency Autism Coordinating Committee (IACC), a group of scientists and advocates charged with guiding federal funding of autism research in the United States.

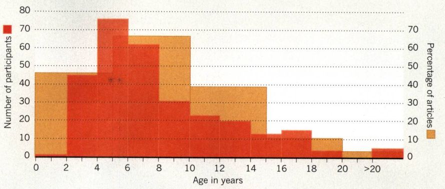
Starting in 2009, the IACC made research on adults with the disorder a strategic priority. That same year, Autism Speaks, a United States–based autism science and advocacy organization, launched a major initiative for adults with the disorder that calls for legislation to increase access to services for these adults.

A 2012 report funded by Autism Speaks estimates that the annual cost of caring for people with autism in the United States is US$137 billion, most of which is for adults. The organization is measuring the economic benefits of providing services to individuals with autism as they transition from adolescence to adulthood.

Individual researchers, too, are turning their attention to adults with autism. For example, Catherine Lord, who developed the gold-standard screening tests for autism, is refining the tools used to diagnose the disorder in adults. Lord is a psychologist who directs the Center for Autism and the Developing Brain, a subsidiary of Weill Cornell Medical College in New York and New York Presbyterian Hospital in White Plains. She has been following 200 young adults since they were diagnosed with autism at 2 years of age. "We have a very good sense of who these kids were when they were little and now what they're like at 20," says Lord. "That's been a huge source of information for us."

Another group, at the University of Wisconsin–Madison, is following more than 400 families with a history of autism. Their study, which started 12 years ago, is focusing on how symptoms of autism change over time and how the changes affect parents and siblings. (A handful of long-term studies were conducted in past decades, but they are difficult to interpret because the diagnostic criteria and techniques for assessing autism have changed over time.)

In that context, Donald Triplett's case raises important questions that we're only beginning to answer. Can children diagnosed with autism expect to live a full life 30, 40 or 50 years from now? Can a subset of children with the disorder outgrow their symptoms, or at least learn to cope



OUT OF SIGHT

Adults are rarely included in studies of behavioural interventions for individuals with autism.

EDWARDS, T. L. *RES. AUT. SPECT. DIS.* **6**, 996–999 (2012).

with them? And, perhaps most importantly, how can we improve the odds that they do?

Most studies show that autism's core deficits, including social and communication problems, persist into adulthood, and that most adults with the disorder remain dependent on family members for support and struggle to fulfil their basic needs such as housing and employment. But work over the past decade or so has offered a more hopeful scenario, partly because many of the children diagnosed today are less severely impaired than those in the past.

In some cases, autism's core symptoms do seem to diminish over time. And a 2009 study of 120 individuals first diagnosed with autism in adulthood found that one-quarter had a college or university degree, and more than 40% held a job or were studying[1]. About half of the participants aged over 23 were living independently.

But the researchers also found higher rates of anxiety and depression in these adults, all of whom had average IQ scores.

According to Julie Taylor, a developmental psychologist at the Vanderbilt Kennedy Center in Nashville, Tennessee, who studies the transition to adulthood for people with autism, only about 10% have independent employment in the community. (The numbers are slightly better — ranging from about 25% to 45% — if they include community-based jobs with support from someone like a job coach or jobs in sheltered workshops.)

The crucial question arises because of this 10%: what determines whether a child with autism is able to graduate from college, hold down a job and live independently?

Some trends are emerging. Faring best are children who score well on standardized intelligence tests, have early verbal skills and display fewer disruptive behaviours such as aggression and hyperactivity. These factors are hard to change, so the challenge to researchers is to find others that can be influenced more readily.

"Knowing about IQ and early language as predictors of better outcomes is all fine and good, but when you have a 12-, 18- or 30-year-old, there's just not a whole lot you can do," says Taylor. So she is beginning to study societal and family-related factors, such as how connected families are, to better design interventions.

The late teen years seem to be particularly critical. One recent study, for example, showed that although autism symptoms generally improve during high school, this improvement slows or ceases after graduation[2]. This is the time when access to specialized services through the school system dries up. A survey of nearly 1,000 young adults (aged 19–23) with autism published in 2011 found that nearly 40% receive no services after graduation[3]. "Parents of kids with autism often liken it to falling off a cliff," says Geraldine Dawson, chief science officer at Autism Speaks.

This abrupt loss of services is particularly true for individuals with average or above-average IQ scores, Dawson says, because they don't qualify for services intended for adults with intellectual disabilities. Higher-functioning adults are three times more likely to have no regular vocational or educational activities than those with lower cognitive function[4]. That the adults who are best equipped to succeed are the ones slipping through the cracks was a real wake-up call, says Taylor.

Whether therapies — vocational, educational, behavioural or drug-based — can help adults with autism meet those expectations is still unclear. Only 32 studies conducted so far of therapies for autism are aimed at adolescents or adults with the disorder (aged 13–30 years), most of which were of poor quality, according to a report published in August 2012 by the US Agency for Healthcare Research and Quality[5]. Compare that with a previous report of 159 studies of treatments that enrolled children 12 years old and under with autism.

> *"Young adults with autism can make enormous progress as they grow up."*

Another 2012 review, by psychologist Al Poling of Western Michigan University in Kalamazoo and his colleagues, found that fewer than 2% of participants in studies of behavioural interventions are aged 20 or older[6].

That's likely to change over the next ten years or so, says Taylor. And that's important because therapies may affect adults and children differently. For example, a small placebo-controlled trial published in 2011 found that adults who take the antidepressant fluoxetine (marketed as Prozac) show improvements in the repetitive behaviours typical of autism. Neither fluoxetine nor other antidepressants have had much success in children with the disorder.

Research on therapies for autism has focused on children because their brains are still developing (see 'In the waiting room', page S14). But there is reason to believe that adult brains also respond to treatment. A mouse study published in 2012 showed that a compound called CTEP reverses many of the neurochemical and behavioural symptoms of fragile X syndrome, an inherited form of intellectual disability that produces an autism-related disorder[7]. The mice were given CTEP at 4–5 weeks of age — the mouse equivalent of human adolescence — suggesting that it may offer benefits after the so-called 'critical window' of brain development has passed.

Another drug, arbaclofen, has undergone phase II clinical testing for fragile X syndrome in children and adults, and is being tested on adults with autism. A new study shows that it alleviates some of the social impairments and behavioural problems that are characteristic of the syndrome.

A third drug, GRN-529, may help people with autism[8]. It significantly improves social behaviour in a mouse model of the disorder compared with controls. "I think the outlook is extremely positive, even for adults," says Randi Hagerman, who led the arbaclofen trial, at the University of California, Davis. "We are pushing the companies hard to do trials for autism with these drugs and they say they will."

Although research on adults with autism is still lagging behind their need for treatment, the fact that autism is now being viewed — and studied — as a lifelong disorder is encouraging.

The 200 children with autism that Lord has followed for the past 20 years are a constant reminder that neural development doesn't cease. "What's been so striking is that young adults with autism can make enormous progress as they grow up," she says. "It's imperative that we look at adult development and how we can support it — from a humanitarian and an economic point of view." ■

Lindsay Borthwick *is a freelance writer based in New Haven, Connecticut.*

1. Hofvander, B. *et al. BMC Psychiatry* **9,** 35 (2009).
2. Taylor, J. L. & Seltzer, M. M. *J. Autism Dev. Disord.* **40,** 1431–1446 (2010).
3. Shattuck, P. T. *et al. Arch. Pediatr. Adolesc. Med.* **165,** 141–146 (2011).
4. Taylor, J. L. & Hodapp, R. M. *Am. J. Intellect. Dev. Disabil.* **117,** 67–79 (2012).
5. Lounds Taylor, J. *Comparative Effectiveness Review* 65 (Agency for Healthcare Research and Quality, 2012).
6. Edwards, T. L. *et al. Res. Aut. Spec. Dis.* **6,** 996–996 (2012).
7. Michalon, A. *et al. Neuron* **12,** 49–56 (2012).
8. Silverman, J. L. *et al. Sci. Transl. Med.* **4,** 131ra51 (2012).

cow
horse
BREAK TIME!

ISTOCKPHOTO.COM

DIAGNOSIS

Redefining autism

Draft diagnostic guidelines are raising concerns that mild forms of the disorder may no longer be recognized.

BY EMILY SINGER

The word 'autism' once invoked the picture of someone with severe problems, serious challenges in communications and rigid behavioural patterns — perhaps someone like Dustin Hoffman's character in the 1988 movie *Rain Man*. These days, the perception of autism tends to be one of a much milder condition — people may know someone with a child with some form of autism, many of whom go to regular schools.

The definition of autism has changed many times since it was first described in the 1940s. It has evolved from so-called childhood schizophrenia in the 1950s and 1960s, to infantile autism in the 1980s, to the broad autism spectrum disorder (ASD) we know today.

The fifth and newest edition of psychiatry's diagnostic bible, the *Diagnostic and Statistical Manual of Mental Disorder* (*DSM-5*), is due to be published in May 2013. Four years in the making, the *DSM-5* has another stab at refining our concept of autism. The process is always closely watched because the guidelines have broad impact: they are used in the United States and many other countries, by clinicians, pharmaceutical companies, researchers and insurance companies to define psychiatric conditions. In this case, controversy surrounds the goal to make the guidelines sensitive enough to diagnose those who have the disorder, and specific enough to exclude those who don't.

CONSTANT FLUX

The two biggest changes in the *DSM-5* are a shift from three classes of symptoms to two, and the collapse of a number of different autism-related disorders, including Asperger's syndrome (which describes people who have some of the social-interaction deficits common in autism but no impairments in language) under the single heading of autism.

The proposal to erase the boundaries between distinct autism-related disorders has raised fears that some children would no longer meet diagnostic criteria under the *DSM-5* definition, making it difficult to get the help they need. "My worry is there will be a substantive change in the way we make diagnoses, and that will complicate the lives of researchers and families," says Fred Volkmar, director of the Yale University Child Study Center in New Haven, Connecticut, and a critic of *DSM-5*.

Supporters of the *DSM-5*'s autism definition counter that the changes are based on a better understanding of the condition's symptoms. What's more, they say, the preliminary studies indicate that few will be excluded. "To the public it seems like the changes are huge and capricious," says Sally Ozonoff, an expert in autism diagnosis at the MIND Institute at the University of California, Davis, who was not involved in the revisions. "But it's based on a very exhaustive review of all the research, and I think they got it right."

Adding fuel to the fire is that unlike previous versions, the *DSM-5* guidelines have been open to public comment as they were developed. The openness of the process has made for heated — and sometimes emotional — reactions. "It hurt like hell that newspapers were accusing us of trying to hurt kids by denying them services," Susan Swedo, chair of the working group charged with revising the guidelines, told a packed lecture hall at an autism conference in Toronto, Canada, in May 2012.

Swedo, a specialist in behavioural paediatrics at the National Institute of Mental Health in Rockville, Maryland, says the new guidelines are very similar to the old ones. "What's really different is there is more guidance for clinicians," she says, such as how different symptoms might look in toddlers or young adults. The goal of expanding these descriptions, she adds, was to make the guidelines more sensitive to autism in groups that are often overlooked, such as girls and minorities.

A CRITERION FOR CRITERIA

One of the aims of the new guidelines was to help standardize diagnosis. A 2011 study by a team led by psychologist and autism researcher Catherine Lord of Weill Cornell Medical

College in New York suggests that although clinicians give similar scores on autism screening tests, the label they attach to that score — such as classical autism or the milder Asperger's syndrome — is hugely variable[1].

To develop the new autism criteria, researchers analysed data from several large data sets that included children with classical autism, autism-related disorders and other diagnoses, such as language delay, intellectual disabilities and attention-deficit hyperactivity disorder. The children had each been screened with two widely used tests: the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-Revised (ADI-R).

To determine what should be included in the criteria, the researchers first ran a statistical analysis designed to isolate the factors responsible for the diagnosis of an autism-related disorder. They then created different versions of the *DSM-5* and tested how accurately each identified children with autism of different ages, gender and language ability. "This allowed us to look at the proportion of children with various clinical diagnoses who would be included or excluded from ASD with the new criteria," says Lord, a member of the *DSM-5* working group and the creator of both the ADOS and ADI-R.

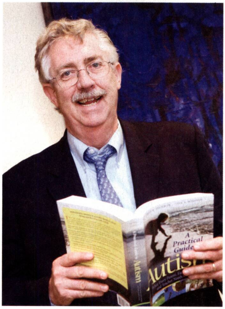
The *DSM-5* collapses the current triad of impairments as described in *DSM-IV* — deficits in social behaviour, language delay, and repetitive or restrictive behaviours — into just two main classes: social communication difficulties, and repetitive or restrictive behaviours. "It was really kind of random whether a particular behaviour was designated as social or non-verbal communication [in the *DSM-IV*]," says Lord. "When we looked at various large data sets, there were two separate factors in ASD." *DSM-5*'s simplification, she says, more clearly reflects the reality of the condition.

Two studies that analysed different sets of data support this change[2,3]. "The conflation of communication and social reciprocity into one category fits very much into our clinical experience," says David Skuse, a developmental neuropsychologist at University College London who led one of the studies. (Skuse is not a member of the *DSM-5* group.)

The more controversial change in the *DSM-5* is to bring together a number of formerly distinct disorders under the single heading of autism. The *DSM-IV* recognized several disorders as being distinct, including classical autism, Asperger's syndrome and pervasive development disorder not otherwise specified (PDD-NOS) — a category that applies to many children who meet some, but not all, of the criteria for autism.

Some autism researchers say there is little scientific basis for defining these as different disorders. "Replicable research findings show that it's almost impossible to discriminate people with autism from Asperger's syndrome if you control for intelligence and language level," says Lord. As for PDD-NOS, "when you look at the research, it's basically less severe autism," she says. "No one has come up with characteristics that pull out that group."

One potential impact of the redefinition that its opponents have cited is that some people with Asperger's syndrome and PDD-NOS will no longer qualify for an autism diagnosis and the support services that come with it. For example, Volkmar and his collaborators reanalysed data collected via questionnaires for the *DSM-IV* field trials[4], linking it to criteria outlined in the *DSM-5*. Their 2012 analysis suggests that only about 60% of 1,000 people diagnosed with autism in 1994 under *DSM-IV* would meet the new diagnostic criteria.



Fred Volkmar is a critic of plans to redefine autism.

COURTESY OF YALE SCHOOL OF MEDICINE

A handful of studies published over the past year support Volkmar's concern[5–8]. "All show a big drop off in numbers of people diagnosed with ASD under the *DSM-5*," says Johnny Matson, a psychologist at Louisiana State University in Baton Rouge and author of some of these studies. "I would rather see a tweak than a major change."

But critics of these studies point out that they are retrospective, meaning they used old data, often from limited questionnaires. Some of the symptoms in the new guidelines, such as sensory deficits, may simply not have been collected in the past. The result, they say, is that individuals who would meet the diagnostic criteria if assessed today would not be considered autistic based on the previously collected information. Retrospective studies are "always going to underestimate sensitivity, because they didn't ask all the questions", says Lord.

The Volkmar study was "fatally flawed because it used data to do the analysis that couldn't answer questions being asked", says Swedo. "Just because you have an idea of how a child met criteria for the *DSM-III* or *DSM-IV* doesn't mean it can be applied to the *DSM-5*."

A new study, published this month in the *American Journal of Psychiatry* by Lord and collaborators, uses a much richer data set to show that the new guidelines are just as sensitive as the DSM-IV and unlikely to exclude many[9].

Experts on both sides agree that the real answer will come from prospective studies, meaning that new data will be collected as participants are simultaneously evaluated with the *DSM-IV* and the *DSM-5*. "We won't know until [results from those] trials are published," says Ozonoff.

Such research is now underway. The research and advocacy organization Autism Speaks is funding a prospective study, as is the American Psychiatric Association (APA), which publishes the *DSM*. Preliminary results from the APA field trials, presented at the International Meeting for Autism Research in Toronto in May 2012, are in line with Lord's study and suggest that the new guidelines are unlikely to exclude many people with an ASD.

Ozonoff, who is broadly supportive of the *DSM-5* guidelines, says one flaw is the creation of a diagnostic category called pragmatic, or social, communication disorder. This label will apply to children who show some of the social communication deficits of autism but not the repetitive and restrictive behaviours. These children are often diagnosed with PDD-NOS. Unpublished data from Skuse's group suggest that about one-third of those people are likely to be shifted to social communication disorder. "It seems very similar to autism spectrum disorder," says Ozonoff. "I wonder about the rationale for creating a new diagnosis unless we have good data that it's different." Just as Asperger's syndrome came and went as a form of autism, the same trajectory may be in store for this latest subcategory.

The only way to resolve the dispute is more research, says Volkmar, who in the 1980s helped evaluate the current criteria in the fourth edition. "Ultimately," he says, "it's not a point for opinion — it's a point for data." ■

Emily Singer *is news editor at SFARI.org, the website of the Simons Foundation Autism Research Initiative in New York City.*

1. Lord, C. *et al. Arch. Gen. Psychiatry* **69,** 306–313 (2012).
2. Mandy, W. P. *et al. J. Am. Acad. Child Adolesc. Psychiatry* **51,** 41–50 (2012).
3. Frazier, T. W. *et al. J. Am. Acad. Child Adolesc. Psychiatry* **51,** 28–40 (2012).
4. McPartland, J. C. *et al. J. Am. Acad. Child Adolesc. Psychiatry* **51,** 368–383 (2012).
5. Worley, J. A. & Matson, J. L. *Autism Spectr. Disord.* **6,** 965–970 (2012).
6. Mandy, W. *et al. Autism Res.* **4,** 121–131 (2011).
7. Taheri, A. & Perry, A. *J. Autism Dev. Disord.* **42,** 1810–1817 (2012).
8. Matson, J. L. *et al. Dev. Neurorehabil.* **15,** 185–190 (2012).
9. Huerta, M. *et al. Am. J. Psychiatry* **169,** 1056–1064 (2012).

UNC ASPIRE RESEARCH PROGRAM

A seven-year-old boy with autism is given oxytocin nasally to see if it alters his sociability.

TREATMENTS

In the waiting room

After years of making do with drugs developed for other conditions, doctors and scientists are eagerly pursuing drugs that target the social symptoms of autism.

BY MICHAEL EISENSTEIN

Why would any parent feed their child worm eggs? In Stewart Johnson's case, it was because he believed he had stumbled across a promising treatment for the symptoms of his teenage son's autism after limited success with other treatments. Although he is not a professional scientist, his bold experiment with porcine whipworm (*Trichuris suis*) ova has culminated in a formal clinical trial spearheaded by a leading autism research centre. However, Johnson's story also tells of the desperation felt by many parents to find something — anything — that might improve the lives of children with autism spectrum disorders (ASDs).

That desperation is borne of the limited success achieved in developing drugs to treat autism. "When I give a lecture on medication, I usually start by saying that I won't be talking about the state of pharmacology in autism, but about the *miserable* state of pharmacology in autism," says Antonio Hardan, a child psychiatrist at Stanford University School of Medicine in California. In fact, no drugs are approved by the US Food and Drug Administration (FDA) for autism *per se*, although a handful of medications approved for other psychiatric indications have proven useful in mitigating some of the symptoms of ASDs. "We have become good at stealing drugs developed for other disorders," jokes Hardan. But even second-hand drugs are better than nothing, given how disruptive these behaviours can be.

> "Talking about immune stuff was considered fringe and funky."

STRIKING A BALANCE

'Irritability' may sound like a minor failing, but the behaviours it encompasses — mood swings, aggression and self-harming — can be debilitating and dangerous to people with autism and their caregivers. That's why irritability is a primary target for pharmaceuticals in ASDs. Two antipsychotic drugs, aripiprazole (marketed as Abilify by Otsuka Pharmaceuticals) and risperidone are now in widespread use in people with ASDs. "These drugs are effective in 60–70% of children and adolescents that have significant irritability," says Christopher J. McDougle, director of the Lurie Center for Autism at Massachusetts General Hospital in Boston. Both drugs, however, come with side effects. Jeremy Veenstra-VanderWeele, a child psychiatrist and researcher at Vanderbilt University Medical Center in Nashville, Tennessee, has determined that aripiprazole and risperidone frequently induce drowsiness and weight gain, and he cautions against prescribing these drugs unless symptoms are severe. "When you're using a medicine in a seven-year-old child that increases their body weight by 10% over the course of a couple of months, that potentially changes the way their metabolism is going to work down the line," he says, noting the chance for increased risk of metabolic syndrome or cardiovascular disease.

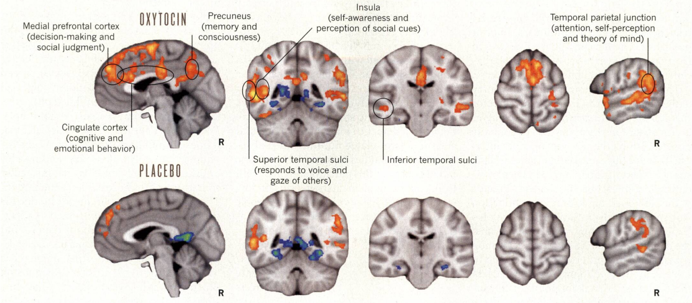
Many clinicians prescribe anti-depressants from the selective-serotonin reuptake inhibitor (SSRI) family, such as fluoxetine (Prozac) and citalopram, to treat ASD patients, despite scant evidence supporting their use. Randomized controlled trials have shown that SSRIs may reduce compulsive, repetitive behaviour in adults with autism, but have no observed effect on these symptoms in children and adolescents. But these trials did not examine social anxiety, a symptom for which some doctors have observed improvements in children with SSRIs. McDougle, who pioneered the study of SSRIs for ASDs while working at Yale University School of Medicine in the 1990s, believes factors related to age-associated brain development may also limit their use. "One reason some children didn't get better was that they couldn't tolerate even low starting doses," he says. "They became much more irritable, agitated and aggressive."

Since the value of both SSRIs and antipsychotics is based largely on the extent to which they enable children with autism to participate in school, caregivers must weigh the potential medical benefits and adverse effects carefully. "If we're putting kids to sleep, even if their aggression is down, we haven't actually facilitated their psycho-educational intervention," says Evdokia Anagnostou, a child

ILANIT GORDON, YALE CHILD STUDY CENTER

ACTIVATING THE AUTISTIC BRAIN

Functional MRI images from a small, preliminary study (7 participants) show that the administration of oxytocin could potentially change brain patterns in a way that would reduce the symptoms of autism.


OXYTOCIN
Medial prefrontal cortex (decision-making and social judgment)
Precuneus (memory and consciousness)
Insula (self-awareness and perception of social cues)
Temporal parietal junction (attention, self-perception and theory of mind)
Cingulate cortex (cognitive and emotional behavior)
R
Superior temporal sulci (responds to voice and gaze of others)
Inferior temporal sulci
R
PLACEBO
R
R

neurologist at the Bloorview Research Institute in Toronto, Canada.

THE HEART OF THE MATTER

No medications target the core symptoms of autism, which include attention deficit, abnormal social behaviour and poor ability to communicate. Accordingly, when a potential treatment option for these symptoms is identified, it grabs the community's attention. For example, extensive research in both animals and humans has identified a prominent role for the hormone oxytocin in behaviours associated with trust, empathy and attachment — emotions commonly impaired in people with ASDs. At least half-a-dozen clinical trials are underway to test whether oxytocin affects these traits in people with ASDs.

The Swiss pharmaceutical company Novartis makes a nasal oxytocin spray, Syntocinon, and this is the primary mode of delivery in development. Veenstra-VanderWeele cautions that the benefits remain unclear. "Kids with ASD are having oxytocin squirted up their noses, and we need to know if that makes any sense," he says. Early studies have shown modest improvements — for example, an increased tendency to make eye contact or elevated activity in social centres of the brain. But most of these investigations have looked at the impact of single doses in small numbers of patients. A new round of larger-scale initiatives should help clarify both the efficacy and the long-term safety of oxytocin use. For example, a treatment network funded by the US National Institutes of Health through its Autism Centers of Excellence programme is set to embark on an ambitious five-year study of 300 children and adolescents to examine whether extended treatment with oxytocin confers measurable improvements in social function.

One of the biggest obstacles confronting clinical studies of oxytocin — and, indeed, any treatment of ASD's core symptoms — is the lack of robust, reproducible measurements of social deficits and improvements in those deficits. "Social behaviour is so complicated, and trying to measure changes in it is really tricky," says McDougle. This is especially true for early-stage experimental drugs, which target neurological pathways with unclear contributions to ASD and in which the therapeutic effects are likely to be weaker than for approved drugs such as risperidone. Furthermore, the inclusivist definition of autism as a spectrum of disorders overlooks very real differences. "Autism is a heterogeneous disorder, and it's very unlikely that you're going to find a compound that's effective in the treatment of social deficits in all kids with autism," says Hardan.

FRAGILE X MARKS THE SPOT

Some of the most exciting progress has emerged from work focused on ASDs with a clearly defined genetic cause. For example, research by Mark Bear, a neuroscientist at the Massachusetts Institute of Technology in Cambridge, has shown that defects in brain signalling mediated by the neurotransmitter glutamate appear to be critical to the pathophysiology of fragile X syndrome (FXS), an inherited disability in which patients often manifest autistic-like social deficits.

FXS is caused by mutations in the gene encoding a protein that regulates the production of other proteins involved in neuron signalling. Bear and colleagues discovered that the protein produced by the mutated gene represses a glutamate receptor called mGluR5. Bear hypothesized that the neurological symptoms of FXS might arise from overactive mGluR5, so his group studied animal models of FXS to test this theory[1]. "We took fragile X mice and engineered them to have 50% of the normal level of mGluR5, with the prediction that this would correct multiple aspects of the disease, and it did," says Bear. "That was a powerful proof of principle."

Since then, numerous drug candidates have been identified that can potentially correct mGluR5-associated deficits — Novartis, Swiss drug giant Roche, and Seaside Therapeutics, based in Cambridge, Massachusetts, and co-founded by Bear, all have agents in clinical trials. Mouse models of FXS have proven useful in testing these candidates, as they directly mimic the genetic and cellular-scale defects seen in FXS, even if they do not always recapitulate the full spectrum of behavioural and social symptoms. Most encouragingly, a collaboration between Bear's team and Lothar Lindemann's research group at Roche has discovered that mGluR5 inhibitors can even repair FXS-associated symptoms in grown mice[2] — suggesting that the window for repairing cognitive damage may be far wider than expected. "That's something we wouldn't have imagined 10 years ago," says Veenstra-VanderWeele, who was not involved in the study.

Novartis has published data from a phase II trial of its mGluR5 inhibitor, AFQ056. The results suggest that the drug may have positive behavioural effects in a subset of adult FXS patients[3], but the size of the cohort that benefited was too small to be conclusive. The company is now recruiting for a larger phase III trial. Seaside has also announced (but not published) encouraging phase II results in treating social symptoms of FXS with arbaclofen, a drug that inhibits glutamate release.

This spate of trials announces the entry of the pharmaceutical industry into a sector that it had long shied away from, owing to issues of complexity and a lack of concrete targets. "I think they are getting more interested — it's a huge market out there, and the science is getting intriguing," says neuroscientist Gerald Fischbach, scientific director of the New York-based Simons Foundation Autism Research Initiative. There is also hope, but not yet supporting evidence, that the deficits seen in FXS may share common roots with some other ASDs. "We need to understand the pathophysiology of each of these models and look for points of convergence, but we should not make assumptions about what those might be," says Veenstra-VanderWeele. Nevertheless, the key role of glutamate signalling in the brain has given some scientists cause for optimism, and several trials are underway to examine the potential benefits of repurposing other known glutamatergic modulators. For example, Anagnostou and colleagues are conducting pilot studies with memantine (marketed as Namenda by Forest Laboratories) and riluzole (marketed as Rilutek by Sanofi-Aventis), which were approved by the FDA for Alzheimer's disease and amyotrophic lateral sclerosis respectively. "This is a very rich area to look for molecular targets," she says.

PICKING UP HINTS

All the same, there is a very real risk that treatments for patients with ASD subsyndromes such as FXS or the neurodevelopmental disorder Rett syndrome will benefit only a small subset of people with autism outside these categories. Some researchers further caution against drawing overly broad connections between these conditions that may be relatively superficial or even misleading. "Everybody says people with those syndromes also exhibit characteristics of autism," says McDougle, "but I can tell you that if you know autism, then you can see that the impaired social relatedness in those disorders is quite different — it's not autism."

Categorization of patients could be improved by the identification of biomarkers — molecular or physiological traits that correlate with specific symptoms or manifestations of ASD. Such markers could steer patient selection for clinical trials, as well as illuminate autism's still-mysterious aetiology. Several candidate genes have been identified, although they have not yet been woven into a coherent model. In April 2012, a large-scale gene sequencing study of 343 families revealed more than 300 potential susceptibility factors — and an unexpectedly strong link to FXS, with many genes encoding targets regulated by fragile X mental retardation protein (FMRP)[4]. "This will still need to be settled with empirical studies in humans," says Bear. In some cases, causative mutations may be exceedingly rare or even limited to certain bloodlines. For example, one recent genetic study identified a link between a defect in a metabolic gene and autism in two related families, and subsequent animal studies suggested that autistic individuals with this mutation might benefit from dietary supplementation[5]. This finding is unlikely to help the vast majority of people with autism, however, and the search continues for broader aspects of ASD pathology.

Some researchers are also beginning to investigate the contributions of inflammatory immune responses and autoimmunity in causing, or at least exacerbating, autism. "Not long ago, talking about immune stuff was considered fringe and funky, and I'm not fringe or funky," says McDougle. "But I think that immune dysfunction is going to prove to be important in a very meaningful subgroup of patients — maybe as many as 10–25%." While the evidence for this correlation remains circumstantial, other scientists are taking the possibility seriously. Stewart Johnson's maverick experiments with whipworm eggs were informed by research data suggesting that the pig parasitic worms, which do not permanently colonize the human gut, can suppress harmful inflammatory responses[6]. Johnson reported improvement in many of his son's behavioural symptoms, and his work was sufficiently persuasive to convince psychiatrist and autism specialist Eric Hollander at the Albert Einstein College of Medicine in New York to embark on a formal clinical trial of this treatment people with ASD .

> *"We have become good at stealing drugs developed for other disorders."*

Feedback from parents has led to other investigations of therapeutic modalities besides the whipworm. For example, input and financial support from parents led Hardan to conduct a clinical trial examining *N*-acetylcysteine, an antioxidant and modulator of glutamate signalling, as a potentially milder alternative to antipsychotic drugs for irritability symptoms[7]. Fischbach notes that parents of children with autism helped call attention to a surprising observation that could offer novel insights into ASD pathology. "One-quarter of the people in our study group of almost 3,000 families report that when their children have a mild fever, their social interaction improves — some start speaking or making more eye contact," he says. "We are trying to formulate controlled trials to understand that phenomenon better."

Not every anecdotal finding pays dividends — Veenstra-VanderWeele cites the prominent example of secretin, a digestive hormone that was described as an effective treatment for core deficits in autism based on a case report[8] of three children published by a research group in 1998, but it ultimately led researchers down a costly rabbit hole. "We figured out three or four years later that it does absolutely nothing," says Veenstra-VanderWeele. "People were spending a lot of time and effort pursuing that instead of something with more evidence behind it." The problem is that parents, desperate for treatment options, are confronted by a staggering amount of ambiguous data — and misinformation — and clinicians must take care to keep parents informed of what has been scientifically confirmed. Anagnostou says her top priority is to steer families pursuing alternative approaches away from known hazards, such as excessively high doses of vitamin B6 or omega-3 fatty acids, or unsafe treatments such as chelation therapy, which entails the use of compounds intended to remove heavy metals from the body. "I take all of the non-evidence-based things that they're interested in and tell them what's dangerous, and then I hope that they will stay in the non-dangerous territory," she says. "But the truth is, we haven't given them that many options."

With luck, this will change as the current round of clinical studies concludes. Clinicians and researchers eagerly await news of the first drug to successfully chip away at autism's core symptoms. However, Veenstra-VanderWeele points out that even a runaway success on the pharmacological front will only be the starting point for effective therapy, enabling people with autism to receive the behavioural and educational training they need to function in society. "You can set the brain up to learn," he says, "but you still need to teach." ■

Michael Eisenstein *is a freelance science writer based in Philadelphia, Pennsylvania.*

1. Dölen, G. *et al. Neuron* **56,** 955–962 (2007).
2. Michalon, A. *et al. Neuron* **74,** 49–56 (2012).
3. Jacquemont, S. *et al. Sci. Transl. Med.* **3,** 64ra1 (2011).
4. Iossifov, I. *et al. Neuron* **74,** 284–299 (2012).
5. Novarino, G. *et al. Science* doi: 10.1126/science.1224631 (2012).
6. Summers, R. W. *et al. Am. J. Gastroenterol.* **98,** 2034–2041 (2003).
7. Hardan, A. Y. *et al. Biol. Psychiatry* **71,** 956–961 (2012).
8. Horvath, K. *et al. J. Assoc. Acad. Minor Phys.* **9,** 9–15 (1998).

PERSPECTIVE

Imaging autism

Several studies in the past two years have claimed that brain scans can diagnose autism, but this assertion is deeply flawed, says **Nicholas Lange**.

A study published recently in the widely read clinical journal *Radiology* correctly identified 36 out of 39 children and adolescents with autism solely by the use of functional magnetic resonance imaging (fMRI)[1]. The researchers reported that as the subjects listened to their parents' voices, those with autism had lower brain activity in the superior temporal gyrus, a brain region involved in language reception, compared with controls.

Although this finding makes sense, we should be cautious about its interpretation. This is not an autism-detecting brain scan. Language deficits are a core and defining feature of the disorder; one doesn't need a brain scan to show this. And there are many people with language difficulties who do not have autism. Therefore, a technology that detects language problems doesn't move the ball towards a differential diagnosis.

To diagnose autism reliably, we need to better understand what goes awry in people with the disorder. Until its solid biological basis is found, any attempt to use brain imaging to diagnose autism will be futile. I agree with my colleagues who have said that any candidate biomarker for any disorder is of no clinical use unless we first establish a stable and biologically valid concept of the illness[2].

> UNTIL ITS SOLID BIOLOGICAL BASIS IS FOUND, ANY ATTEMPTS TO USE BRAIN IMAGING TO DIAGNOSE AUTISM WILL BE FUTILE

During the 70 years since psychiatrist Leo Kanner originally observed and named autism, its diagnostic criteria (see 'Redefining autism', page S12) have not included any biological measurements. Elevated blood level of the neurotransmitter serotonin is its only candidate biological marker to date, but it is hardly specific to autism. For instance, tumours in the gastrointestinal tract and recreational drug use also increase blood serotonin.

Given this absence of a simple physical test for autism, it's tempting to think that a 5- or 10-minute brain scan can do the trick, by quickly and accurately identifying something distinctive in the brain of an individual with autism. But we don't yet know its biological underpinnings, and we cannot expect brain imaging to diagnose autism without that *a priori* biology. To believe otherwise is to place false confidence in technology[3], and potentially to mislead or even harm affected individuals, their parents, families and communities.

Brain imaging does of course have a place in autism study. While it has little if any diagnostic utility at present, this technology has already helped us understand the disorder. Volumetric MRI has shown us that about one in five children with autism has an early abnormal enlargement of the brain that tends to stabilize during the first 18 months of life[4,5]. Functional MRI has helped us learn where people with autism focus during social interaction, or when watching films with intense social content. Positron emission tomography (PET) has begun to show us differences in the distribution of serotonin and dopamine receptors — and of the evolving roles of serotonin throughout life — in the brains of people with autism compared with typical controls. Indeed, if the underlying autism pathology puzzle is solved, PET may provide some guidance as a marker of drug effectiveness. Similarly, magnetic resonance spectroscopy may eventually prove helpful in treatment studies. For instance, it may help quantify differences between the neurotransmitters glutamate and gamma-aminobutyric acid (GABA) in people with autism and those with schizophrenia.

A correct diagnosis of autism often benefits an individual. A false diagnosis can do harm. It is therefore essential — but unfortunately rare — to publish the positive predictive value of a proposed test (the proportion of correct diagnoses among all diagnoses). One of my own autism studies offered an MRI-based classification algorithm (not diagnostic) with the highest overall ability to date, having 90% positive predictive value[6]. But, like other such studies, mine was limited by small sample and effect sizes and requires external replication.

What individuals with autism and their parents urgently need is for us to conduct large, long-term, multicentre studies to identify physical features unique to the brains of individuals with autism. We should pay particular attention to differences in genetics, physiology, immune challenge profile, neurochemistry and brain connectivity between individuals with autism, those with other developmental disorders, and typically developing individuals. These studies need to enrol patients across the entire spectrum, and should use existing reference imaging databases of typical brain development.

No one has conducted a study of this magnitude or potential authority, and reductions in research funding make it a tough sell. Yet because autism is an individually heterogeneous disorder, it is only in this way that we can achieve the goals of a future 'stratified psychiatry'[2]. This term refers to a newly proposed system for all psychiatric disorders in which homogeneous subgroups of individuals are defined by dimensional combinations of biological and behavioural features — such as an altered neurotransmitter pattern and a history of major depression — rather than by categorical criteria.

Autism prevalence worldwide is rising rapidly. In the United States, prevalence jumped by 78% between 2002 and 2008 (ref. 7). Larger numbers and solid biology may mean, however, that we can power our studies with enough participants to find reliable, sensitive and specific markers of the disorder with clinically acceptable predictive value, and improve the lives of people with autism.

Nicholas Lange *is a biostatistician at Harvard Medical School in Boston, Massachusetts.*
email: nlange@hms.harvard.edu

1. Lai, G. *et al. Radiology* **260,** 521–530 (2011).
2. Kapur, S., Phillips, A. G. & Insel, T. R. *Mol. Psychiatry* doi: 10.1038/mp.2012.105 (2012).
3. Racine, E. *et al. Nature Rev. Neurosci.* **6,** 159–164 (2005).
4. Nordahl, C. W. *et al. Proc. Natl Acad. Sci. USA* **108,** 20195–20200 (2011).
5. Lainhart, J. E. & Lange, N. *J. Am. Med. Assoc.* **306,** 2031–2032 (2011).
6. Lange, N. *et al. Autism Res.* **3,** 350–358 (2010).
7. http://www.cdc.gov/ncbddd/autism/addm.html.

RICHARD JONES/REX FEATURES

CULTURE

Diverse diagnostics

The study of autism around the globe must account for a variety of behavioural norms in different societies.

BY SARAH DEWEERDT

In rural South Africa, young children may look at adults' faces while having a conversation, but they don't usually make direct eye contact because it is considered disrespectful. Yet a lack of eye contact is a hallmark of social deficits in people with autism, and as such it is something Western clinicians look for when diagnosing the disorder.

There are other examples of children's behaviour — such as finger pointing to draw attention to something, or conversing with adults as if they are peers — that are commonplace in the West and included in tests of autism.

"Most autism research originates in the West, and we have a particular view of what autism is, a particular view about how children behave and interact with adults," says Courtenay Norbury at Royal Holloway, University of London, who worked with children with autism from ethnically diverse backgrounds in east London[1]. "Other cultures might have very different expectations of how children behave."

This viewpoint makes it challenging to use behavioural diagnostic tests for autism in places where the disorder may look — and even be — different from in the West. But with growing interest in autism's true prevalence worldwide and the need for autism services in poor countries, researchers are grappling with the best ways to objectively diagnose the disorder.

Parent support groups for autism exist in more than 100 countries. "We know that autism is diagnosable and observable across cultures," says Mayada Elsabbagh, a researcher at McGill University in Montreal, Canada, who is leading a group within the International Society for Autism Research on cross-cultural issues. "But the exact details of how different cultures or settings modify autism is unknown," she says.

For many years, the relevance of culture to autism was ignored. Some researchers believed that autism was intrinsically linked to modernity and Westernization, and was rare in other cultures. Others assumed that because autism is a neurobiological disorder, its expression should be the same everywhere[2].

But many researchers are beginning to take a subtler point of view. "While autism itself, the neuropathology of it, may not be culturally determined, our interpretation of those behaviours and our response to those behaviours is," says David Mandell, associate director of the Center for Autism Research at the Children's Hospital of Philadelphia in Pennsylvania.

DIAGNOSTIC DIFFERENCES

Before they can assess autism worldwide, researchers must measure how often various behaviours occur in different cultures and establish the norm, says Charles Zaroff, a psychology researcher at the University of Macau in China. They will also need to work with parents of children with autism to identify how it might manifest in a given culture.

In many Asian cultures, for example, children are expected to express respect for their elders through their language and behaviour, but these aspects of social interaction can be tricky for children with autism to master. Such difficulties would hardly be noticed in much of the West. "Lacking that deference would appear completely appropriate in the United States, but that lack of attention to the strata in society based on age would appear very abnormal in places like China," Zaroff says.

Because the most widely used screening tests for autism were developed in the United States and the United Kingdom, researchers are finding that they have to adapt the tests to identify autism in other countries. For example, part of the Autism Diagnostic Observation Schedule (ADOS), one of the gold-standard diagnostic instruments for the disorder, involves observing a child having a pretend birthday party

— singing 'Happy Birthday', cutting and distributing slices of cake, and so on. But in rural areas of South Africa, birthdays often aren't celebrated, so even typically developing children might be unfamiliar with this ritual.

So, for a study of early autism diagnosis in KwaZulu-Natal province, researchers developed an alternative scenario of shared excitement, involving a traditional African song. "It's finding the intention of what you're trying to elicit, and then finding an alternative," says Amy Wetherby, director of the Autism Institute in the College of Medicine at Florida State University in Tallahassee, who led the work.

Wetherby developed a list of 22 early signs of autism based on studies of several hundred children in Florida. Unpublished data from 19 children show that many of the same red flags differentiate children with autism from their typically developing peers in South Africa as well, says Wetherby.

"The amazing part to me is the unusual gestures," says Wetherby. For example, rather than pointing or looking together at an object, children with autism in the United States may communicate by taking an adult's hand and moving it to the object. "That's an early sign of autism, and we see that [in South Africa] as well."

One reason the patterns of autism symptoms appear to be similar across cultures may be that the participants in this study are only 18–36 months of age. "The earlier we go," Wetherby says, "the more similarities we will see."

More culturally specific or environmental symptoms of autism may emerge as children grow up. For example, according to one recent study, 5–12-year-olds with autism in the United States are more sensitive to sights and sounds than are children with autism in Israel — although the authors note that genetic, cultural and environmental factors might be an influence, as well as the parents' reporting of their child's behaviour[3].

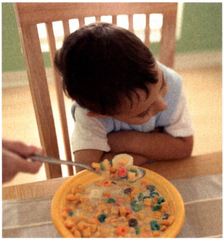
SUBTLE SPEECH

If cultural differences emerge later in childhood, this could further complicate the diagnosis of autism. The disorder tends to be diagnosed later outside the United States and Western Europe, in part because of a lack of awareness of both developmental norms and autism. A parent who notices that a child is withdrawn or has a language delay may not recognize that as a symptom of autism.

In-depth interviews in Goa, India, show that the parents there aren't attuned to early social and communication milestones, and they usually become alarmed only when a child starts preschool and has trouble connecting with peers[4]. "What really concerns the parents initially is, 'he's not fitting in with everybody else'," says Gauri Divan, a paediatrician working with the child health organization Sangath in Goa.

In some cultures, parents may notice symptoms that are not typically associated with autism. Among Latino migrant workers in Florida, for example, "the first complaint seems to be that the child is a picky eater," says Roy Richard Grinker, a George Washington University anthropologist in Washington DC, who is collaborating with Wetherby on a study of autism in this community. "But then, if you go into more detail, you start to see that these children the mothers are describing are probably going to fall on the autism spectrum." Grinker speculates that these mothers are particularly aware of eating habits because they are poor and food is scarce.

Some evidence suggests that doctors need to be trained to spot the signs of autism from oblique comments made by parents. For example, Mandell says, white parents in the United States often emphasize a child's lack of communication by saying, 'my child doesn't respond when I call his name', while black parents tend to use phrases like 'my child won't mind me'. Doctors may be less apt to consider a diagnosis of autism when they think a parent is describing a disobedient child rather than a socially impaired one — possibly helping to explain why autism is diagnosed less frequently among black children.


FUSE/ GETTY IMAGES

Picky eating might point to autism, but a parent noticing a child's eating habits is cultural.

RAISING RATES

If parents in different cultures developed the same sense of autism awareness as in the West, research suggests, autism prevalence around the world might look no different to — or may be even higher than — in the United States or the United Kingdom.

Perhaps the most dramatic demonstration of this is a study of more than 55,000 children in South Korea, which estimated autism prevalence at 2.64% (ref. 5). That's more than twice the autism prevalence in the United States estimated by the US Centers of Disease Control and Prevention, and more than 50 times higher than the South Korean government's figure for autism prevalence of 0.046% (ref. 6). One reason for that higher estimate may be that the researchers screened children in the general population for autism symptoms, rather than recruiting them only from clinics for autism and other developmental disorders.

In South Korea, researchers suggest that one reason for the underdiagnosis in autism may be that the stigma attached to the disorder is particularly strong in that country[7]. The diagnosis of a Korean child with autism diminishes the marriage prospects of siblings, and it can even affect his or her parents' careers. Parents often prefer that their child be labelled as having 'reactive attachment disorder', or 'lack of love' as it's known in Korean, a diagnosis that affects the mother's reputation.

Still, about two-thirds of the children the South Korean study identified as having autism attended mainstream schools and were not receiving any autism-related services. This widespread mainstreaming raises the question of whether a Western-defined autism diagnosis is meaningful if children are able to function reasonably well in their cultural context. "That's an interesting issue," says Norbury, who questions "whether we should be worried about these kids, and whether we should be making families worried about them, if there wasn't any kind of worry before."

Other children identified in the study may be what Koreans would call a 'border child,' a new term that is emerging to describe some who would probably be diagnosed with autism in the West[7]. "This is a child who is high-functioning enough to be in a mainstream school, but who has significant social impairment," Grinker says. Parents prefer this label, he says, because it implies that the child is only impaired socially, not intellectually, and that the condition is temporary.

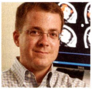
In stigmatizing autism, South Korea is not unique. But Elsabbagh points out that attitudes towards a disorder often change as more resources and services become available. That, she says, is one powerful argument for more cross-cultural research on autism.

Another is that it will advance understanding of the biology of autism. "We've constrained our participant pools to those of European ancestry, and we have also not considered very thoroughly some of the cultural determinants that may shape autism in different ways," says Elsabbagh. "Taking a more global perspective would allow us to see that underlying commonality much more easily." ■

Sarah DeWeerdt *is a science writer based in Seattle, Washington.*

1. Norbury, C. F. & Sparks, A. *Dev. Psychol.* (published online 5 March 2012).
2. Daley, T. C. *Trans. Psychiatry* **39,** 531–550 (2002).
3. Caron, K. G. *et al. Am. J. Occup. Ther.* **66,** e77–e80 (2012).
4. Divan, G. *et al. Autism Res.* **5,** 190–200 (2012).
5. Kim, Y. S. *et al. Am. J. Psychiatry* **168,** 904–912 (2011).
6. Kang-Yi, C. D. *et al. J. Autism Dev. Disord,* (published online 22 June 2012).
7. Grinker, R. *et al. Autism Res.* **5,** 201–210 (2012).

PERSPECTIVE



Kevin Pelphrey

Brain scans need a rethink

Head movement can bias brain imaging results, undermining a leading theory on the cause of autism, say **Ben Deen** and **Kevin Pelphrey**.

One of the most popular and widely accepted theories on the cause of autism spectrum disorders attributes the condition to disrupted connectivity between different regions of the brain. This 'connectivity hypothesis' claims that the social and cognitive abnormalities in people with autism can be explained by a dearth of connections between distant regions of the brain[1]. Some flavours of this theory also predict more connections between nearby brain regions.

Recent studies, however, have found that when a person moves their head while undergoing functional magnetic resonance imaging (fMRI) — a method that maps how different neuroanatomical structures of the brain interact in real time, its functional connectivity — it looks like the neural activity observed in autism. That's a sobering discovery: it means that a major source of evidence for a leading hypothesis on autism, and one that several research teams have pursued for years, may arise from an artefact.

A MAJOR SOURCE OF EVIDENCE FOR A LEADING HYPOTHESIS ARISES FROM AN ARTEFACT

Many studies have investigated functional connectivity in the brains of people with autism, and most have reported evidence supporting the connectivity hypothesis. These findings are consistent with results from some animal models of autism, and from studies using diffusion tensor imaging, which measures the bundles of fibres connecting parts of the brain.

But three studies published in 2012 have come to the same conclusion: head motion leads to systematic biases in fMRI-based analyses of functional connectivity[2–4]. Specifically, motion makes it appear as if long-range connections are weaker than they really are, and that short-range connections are stronger than they really are.

This bias affects all functional connectivity analyses, but it is particularly insidious for studies of autism. That's because it would lead to precisely the patterns that have been observed in fMRI scans of children with autism, and because children with autism typically move more than unaffected children do.

How can autism researchers overcome this bias? One approach would be to define a measure of head motion in each participant over the course of a scan — for instance, to compute the displacement of the head between consecutive time points, and average this over time. Researchers can then check that the autism and control groups are well matched on this measure, or include this value as a nuisance variable in regression analyses.

The matching, however, would need to be precise: as one of the new studies showed, even a difference as small as 0.004 millimetre in average head motion across groups of patients can lead to significant differences in correlation strengths[4].

Furthermore, it is likely that motion artefacts can persist even when groups are matched on mean head motion. First, there is some evidence that head motion relates to functional connectivity measures in a nonlinear fashion[2,4]. If that's true, it would not be sufficient to account for only linear effects of motion. Even if head motion doesn't significantly differ across groups, a nonlinear function of head motion could.

Second, a given estimate of average motion can correspond to rather different scenarios — a few isolated but large movements, or constant small movements — which would have different effects on fMRI signals and on measures of functional connectivity. For instance, it has been shown that large, jerky movements lead to brief spikes in fMRI signal strength.

Based on this observation, cognitive neuroscientist Steven Petersen and his colleagues at Washington University in St Louis, Missouri, propose a strategy for mitigating these head-motion artefacts[2]. They recommend removing periods of high motion. They have shown that this technique, which they call 'scrubbing', corrects at least some of the spurious correlations caused by head motion.

This approach is promising but many questions remain. For instance, the extent of motion-induced correlations even after scrubbing is not well understood. Moreover, the optimal motion threshold for removing periods has not been studied in detail. Still, performing scrubbing in addition to group matching on average motion estimates and other standard noise-reduction methods represents current best practice in functional connectivity research.

So far, only a few research groups have used scrubbing or similar techniques in fMRI studies on autism[5,6]. Given the importance of understanding whether previous results on functional connectivity in autism are real or merely motion-related artefacts, we hope that the authors of fMRI functional connectivity studies in autism (and particularly those reporting results consistent with the connectivity hypothesis) will reanalyse their data using these techniques.

Scrubbing can be easily implemented using freely available software such as Artifact Detection Tools or by implementing custom modifications to the programs. Reanalysing data should be facilitated by tools provided by the US National Database for Autism Research, part of the National Institutes of Health. In addition, a trove of information has recently become publicly available in the Autism Brain Imaging Data Exchange — a collection of resting-state fMRI imaging data sets from 539 individuals with autism and 573 controls. Revisiting fMRI studies with these approaches would help establish whether there really is a connectivity deficit in the brains of people with autism.

Ben Deen *is a graduate student in brain and cognitive sciences at the Massachusetts Institute of Technology in Cambridge.* **Kevin Pelphrey** *is the Harris Associate Professor in the Child Study Center at Yale University in New Haven, Connecticut.*
email: kevin.pelphrey@yale.edu

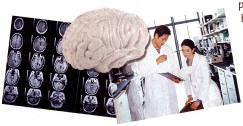
1. Just, M. A. *et al. Brain* **127,** 1811–1821 (2004).
2. Power, J. D. *et al. Neuroimage* **59,** 2142–2154 (2012).
3. Satterthwaite, T. D. *et al. Neuroimage* **60,** 623–632 (2012).
4. Van Dijk, K. R. A. *et al. Neuroimage* **59,** 431–438 (2012).
5. Kennedy, D. P. *et al. Neuroimage* **39,** 1877–1885 (2008).
6. Deen, B. *et al. International Meeting for Autism Research* (Philadelphia, 2010).

SPONSOR FEATURE

Autism Speaks is the world's leading autism science and advocacy organization. We are dedicated to funding research into the causes, prevention, treatment and early diagnosis of autism; increasing awareness of autism and advocating for the needs of individuals with autism and their families. Autism Speaks was founded in 2005 by Suzanne and Bob Wright, grandparents of a child with autism, and has the goal of changing the future for all who struggle with ASD.

Autism Speaks has committed more than $180 million to scientific research spanning more than 750 studies, fellowships and scientific initiatives. All of these address one or more of our key scientific aims:

- Discover and deepen understanding of autism's genetic and environmental risk factors;
- Develop improved methods for early detection and delivery of effective early intervention;
- Identify and decipher biological mechanisms that can become the target of improved treatments;
- Develop and evaluate new treatments, support services and other interventions that improve function and quality of life across the lifespan;
- Enhance access to effective, evidence-based care in all communities, including underserved groups, around the world.



Autism Speaks resources that promote scientific discovery and clinical practice guidelines include the Autism Genetic Resource Exchange, Autism Genome Project, Autism Tissue Program and the Autism Speaks Autism Treatment Network (ATN). Autism Speaks ATN is the first network of medical centers dedicated to comprehensive care of individuals with ASD and emphasizes treatment of associated medical conditions.

Critical to the evolution of Autism Speaks is an increasing emphasis on translational research that accelerates the pace at which basic scientific discoveries become new and effective ways of diagnosing and treating ASD. This includes "bench to bedside" investigations that move the most promising medicines and other interventions from the laboratory into clinical trials – always with the goal of improving the lives of individuals on the autism spectrum.

We invite clinical, basic and applied researchers, as well as pre- and post-doctoral trainees to explore Autism Speaks funding opportunities on our website:

www.autismspeaks.org/science/grants-program

AUTISM SPEAKS
It's time to listen.
www.AutismSpeaks.org

SPONSOR RETAINS SOLE RESPONSIBILITY FOR CONTENT

SPONSOR FEATURE

The Autism Science Foundation provides funding directly to scientists who are conducting cutting-edge autism research to discover the causes of autism and to develop better treatments. ASF also provides information about autism to the general public and supports the needs of individuals with autism and their families.

ASF was founded in 2009 by parents and scientists working together to improve the lives of people with autism. We believe that outstanding research is the greatest gift we can offer our families. We are committed to understanding the causes of autism and to finding a cure for autism, but until a cure is found, our best weapons against autism are early intervention, new treatments and evidence-based services and supports for people with autism of all ages. We must continue to invest in research to discover new and better treatments for the individuals with autism who are here today, as well as focusing on the future.

AUTISM SCIENCE FOUNDATION

The Autism Science Foundation is committed to supporting young investigators with cutting-edge ideas who need seed money to get their research off the ground. In 2012, we increased our funding for pre- and post-doctoral research fellowships by 50% and we expect this trend to continue in 2013. The government and other funders have decreased funding for training grants, but we are determined to help young scientists gather the initial data they need to attract major funding from the National Institutes of Health (NIH), all while encouraging the best and brightest researchers to dedicate their careers to autism research.

In August of 2011, the nonprofit rating agency Guidestar named the Autism Science Foundation the number one startup nonprofit in the national "Disabilities" category. Guidestar called ASF "a shining star to those interested in real science and evidence-based interventions". Guidestar also wrote that "ASF has made a great contribution to empirically valid information on autism spectrum disorders", that "this group funds outstanding autism research that is totally science based", that this group has "strong leadership and operates with the highest level of integrity", and that "ASF has extremely low overhead; every dollar possible is put toward autism research."

Learn more about the Autism Science Foundation at

autismsciencefoundation.org

SPONSOR RETAINS SOLE RESPONSIBILITY FOR CONTENT

Simons Foundation Autism Research Initiative

www.sfari.org

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

SFARI's goal is to support research that will lead to a better understanding of autism spectrum disorders and improve the quality of life for individuals on the autism spectrum

Autism Spectrum Disorder (ASD) is a prevalent developmental disorder. It is also a multifaceted, complex disorder. Such complexity is not unique to autism, of course, but it presents great challenges.

SFARI has approached these challenges by recruiting and supporting scientists to conduct studies of gene discovery and mechanisms of gene action, and to explore how these actions might affect cognition and behavior. In addition to this bottom-up approach, we support top-down research, beginning with behavior and searching for underlying neural mechanisms that may be common to those with ASD. The magnitude of this effort is reflected in our budget of $60 million per year.

Simons Simplex Collection

The Simons Simplex Collection (SSC) was established to optimize the discovery of *de novo* mutations that likely arise in the germ cells of one parent[1]. More than 2,700 simplex families, nearly all of which consist of one child with autism, two unaffected parents and at least one unaffected sibling were carefully evaluated at one of 13 university clinics across North America. Efforts were made to ensure reliability in administration and scoring of the numerous assessments. Inclusion of an unaffected sibling in the SSC has proven invaluable in identifying *de novo* mutations.

To date, more than 100 *de novo* genetic variants that increase risk for ASD have been identified. Each variant is rare, but together they may account for more than 25 percent of all individuals on the spectrum. This advance in knowledge is remarkable, as no such variants were recognized a few years ago. These new gene targets in turn offer opportunities for the discovery of variants that may contribute to risk, or may modify the impact of more penetrant variants. They also offer opportunities for the identification of relevant metabolic pathways, brain circuits and therapeutic targets. Biospecimens (whole blood DNA, serum, and lymphoblastoid cell lines) and phenotypic data from all participating family members are available to researchers via SFARI Base (see below).

Recently, the SSC has joined forces with the Interactive Autism Network (IAN) to create the Simons Simplex Community at IAN. This collaboration will facilitate follow-up studies and maintain the community of study families so that the SSC will remain a valuable resource for years to come.

Simons Variation in Individuals Project

In a novel 'genetics first' approach, 100 individuals with a deletion and 100 with a duplication of chromosomal segment 16p11.2, the most recurrent variant observed in the SSC, are being evaluated clinically and by structural and functional brain imaging with uniform protocols[2].

Participants are recruited in various ways, including referral by clinical genetic centers or testing laboratories, by web-based networks and by self-referral of families who learn about the Simons Variation in Individuals Project (Simons VIP). The Simons VIP has worked to develop a community of 16p11.2 families (Simons VIP Connect) and to provide access to cutting-edge clinical and research information.

Careful analysis of genetically defined autism subtypes will allow detailed phenotypic comparisons within and among these groups. DNA is extracted from blood samples and fibroblasts are cultured and cryopreserved from skin biopsy samples. As with the SSC, phenotypic data and biospecimens are available to approved researchers.

SFARI Base

SFARI Base, developed with Prometheus Research, LLC, is a web-based information system that supports data acquisition, curation and distribution for scientific studies and related resources, such as the SSC and the Simons VIP.

SFARI Base enables scientists to explore de-identified clinical data, place orders for biospecimens and obtain data produced by genetic analysis. Researchers return their results to SFARI Base. A future goal is for the system to store or link to results from DNA sequencing, electrophysiological, neuroimaging, cognitive and behavioral studies.

SFARI Gene

SFARI Gene, developed in collaboration with Mindspec Inc., is a database containing genes implicated in autism. SFARI Gene is designed to permit quick entrée into the literature and to help researchers keep up with the identification and understanding of the myriad genetic risk factors now emerging[3]. This online resource offers five modules, each offering a different way of accessing information about autism risk genes.

Human Gene Module lists more than 284 genes implicated in autism, with annotations and links to published papers. Each gene has a separate page that features intuitive, tab-based cross access to information about that gene residing in other modules.

Gene Scoring Module offers a critical evaluation of the evidence for each gene's association with ASD. Six expert geneticists collaborated on this effort, which will be continually updated with input from the scientific community.

Copy Number Variant Module provides exhaustive detail on all copy number variants reported in individuals with ASD.

Protein Interaction Module curates protein-protein and protein-nucleic acid interactions reported in the literature for the genes in the Human Gene Module.

Animal Model Module lists more than 288 lines of mice that carry mutations in genes described in the Human Gene Module. The animal model list is accompanied by extensive information about each line of mice, including information on phenotypes of greatest relevance to ASD.

In addition, SFARI has placed some animal models referenced in SFARI Gene bearing the most convincing human mutations at Jackson Labs for distribution to the scientific community. Changes in social behaviors have been documented in several strains, as have changes in spatial memory, working memory, repetitive movements, and region-specific expression of candidate genes. Noteworthy are studies that have drawn attention to a subset of GABAergic neurons in the cerebral cortex and to catecholamine neurons in the brain stem. Analogs of human mutations are also under investigation in flies, worms and zebrafish.

Investigators

One of the most remarkable and encouraging happenings in recent years is the influx of talented investigators from many areas of the life sciences to the study of autism. More than 120 principal investigators are now supported by SFARI, many of whom were not working on autism five years ago. These individuals were attracted to work on targeted projects such as gene discovery, and by calls for investigator-initiated, peer-reviewed grants. A collaborative community has been established through frequent workshops, an annual meeting, a policy of early and open data sharing, and by ongoing communication with the SFARI science staff (Marta Benedetti, Marian Carlson, Wendy Chung, Alice Luo Clayton, Alan Packer and Julia Sommer). Collaboration is also enabled by participation in lively, secure discussion forums supported by our newly re-designed website, SFARI.org.

Next steps

Efforts to define the genetic landscape will continue, with even more emphasis on gene networks. Two or three emerging networks involve synapses and the immune system.

It is likely that ASD and other neuropsychiatric disorders involve altered function at synapses. In fact, many genetic risk factors point to proteins that alter synapse formation, synaptic function, and/or synaptic plasticity. Changes in the balance between synaptic excitation and inhibition have been documented in several regions of the brain. We must now understand how large populations of nerve cells form relevant neural circuits and the computations they perform. Ultimately, such measures may provide sensitive assays and more relevant biomarkers.

Clarification of the genetic landscape and the creation of animal models will shed light on environmental factors that may play a role in ASD. Epigenetic modifications of gene expression, some of which involve experience-dependent DNA methylation, have been reported. Also, activation of innate and adaptive immune systems have been documented in some mothers of children with ASD.

Finally, additional human subjects research is essential and eminently feasible. A 'social brain network' has been revealed by fMRI during the presentation of socially relevant stimuli such as biological motion, face processing, and theory-of-mind tasks. Biological motion has proven to be an important assay for diagnosis in young children at risk for ASD. Clinical trials of agents that modify the action of GABA, glutamate and of oxytocin, and attempts to replicate the beneficial effects of fever are under way. They are directed at the core deficits of autism: social cognition, restricted interests and repetitive behaviors. Other studies have focused on mTOR and UBE3A pathways.

We believe that lessons learned from studies of ASD will in turn shed light on other neuropsychiatric disorders and on the social brain, one of the greatest mysteries in all of neuroscience.

References

1. Fischbach G.D. and C. Lord *Neuron* **68,** 192-195 (2010)
2. Simons VIP Consortium *Neuron* **73,** 1063-1067 (2012)
3. Banerjee-Basu S. and A. Packer Dis. *Model Mech.* **3,** 133-135 (2010)

Authors

Gerald D. Fischbach
Director, SFARI

John E. Spiro
Deputy Director, SFARI

nature

http://www.nature.com/nature/outlook/autism_2012

RESEARCH

NEW ONLINE

Papers published this week at Nature.com/nature

NEWS & VIEWS A nanotube device that produces true three-dimensional holgrams **p.47**

ARTICLE Gene variation mapped in more than a thousand individuals **p.56**

LETTER The shifting shape of a river system in the wake of climate change **p.87**

PUBLISHED 28 OCTOBER 2012

MEDICINE

Herpes simplex pathogenesis

Childhood herpes simplex encephalitis (HSE) is a rare but severe disease in which the usually innocuous herpes simplex virus 1 infects the central nervous system. Fabien Lafaille *et al.* report the development of an *in vitro* model for HSE. They derived induced pluripotent stem cells (iPSCs) from patients deficient in TLR3 and UNC-93B. Neurons and immature oligodendrocytes from these iPSCs are susceptible to infection by the virus and fail to induce the normal innate immune response. Undifferentiated iPSCs, neural stem cells and astrocytes, however, are not susceptible. This study shows the underlying pathogenesis of HSE in children with TLR3 and UNC-93B deficiencies.

Impaired intrinsic immunity to HSV-1 in human iPSC-derived TLR3-deficient CNS cells
F G Lafaille *et al.*
doi:10.1038/nature11583

IMMUNOLOGY

Accelerated disassembly of IgE–receptor complexes by a disruptive macromolecular inhibitor
B Kim *et al.*
doi:10.1038/nature11546

NEUROSCIENCE

Synaptic amplification by dendritic spines enhances input cooperativity
M T Harnett *et al.*
doi:10.1038/nature11554

A map of visual space in the primate entorhinal cortex
N J Killian *et al.*
doi:10.1038/nature11587

Layered reward signalling through octopamine and dopamine in *Drosophila*
C J Burke *et al.*
doi:10.1038/nature11614

PUBLISHED 31 OCTOBER 2012

ASTRONOMY

Super-bright supernovae

Superluminous supernovae — stellar explosions ten times brighter than type Ia supernovae — were discovered only a few years ago, and are rare in the nearby Universe. A search for high-redshift supernovae has now identified two superluminous supernovae at redshifts of 2.05 and 3.90, both with slowly evolving light curves. These discoveries imply a production rate of superluminous supernovae at these redshifts at least ten times greater than that predicted in the nearby Universe. In addition, the search technology used has the potential to reveal superluminous supernovae at even greater redshifts, at which point it would become possible for astronomers to witness the deaths of the first generation of stars to form after the Big Bang.

Superluminous supernovae at redshifts of 2.05 and 3.90
J Cooke *et al.*
doi:10.1038/nature11521

Cosmic explosions in the young Universe
S J Smartt
doi:10.1038/nature11643

CANCER

Melanoma risk linked to red hair

Individuals with a 'redhead' phenotype — typically pale skin, red hair and an inability to tan — often carry a polymorphism in the gene encoding the melanocortin 1 receptor (*MC1R*) that reduces the receptor's ability to stimulate the production of the black/brown pigment eumelanin from the red/yellow pigment pheomelanin. David Fisher and colleagues report that inactivation of *Mc1r* in mice promotes melanoma formation in the presence of $\mathrm{BRAF}^{\mathrm{V600E}}$, the most common melanoma oncoprotein, independently of exposure to ultraviolet radiation. Because abrogation of all pigment production in the mice abolishes the effects, they conclude that it is pheomelanin synthesis *per se* that promotes melanoma formation, through an increase in oxidative damage. In practical terms, this suggests that further protective strategies, in addition to avoiding sunlight, could be of benefit in at-risk individuals.

An ultraviolet-radiation-independent pathway to melanoma carcinogenesis in the red hair/fair skin background
D Mitra *et al.*
doi:10.1038/nature11624

Complexion matters
M Fukunaga-Kalabis & M Herlyn
doi:10.1038/nature11641

EVOLUTION

How birds got where they are

There are almost 10,000 bird species alive today. Walter Jetz *et al.* now analyse the evolutionary relationships between all these bird populations in time and space. The resultant evolutionary tree reveals that the rate of diversification has increased in the past 50 million years, suggesting that the 'bird niche' is not yet full. Recent evolutionary success is not evenly distributed, with songbirds, waterfowl, gulls and woodpeckers among the winners. Despite greater diversity near the Equator, diversification rates are particularly low in Africa, South-East Asia and Australia.

The global diversity of birds in space and time
W Jetz *et al.*
doi:10.1038/nature11631

Forum: Birds of a feather
R E Ricklefs & M Pagel
doi:10.1038/nature11642

PHYSIOLOGY

Endothelial cell expression of haemoglobin α regulates nitric oxide signalling
A C Straub *et al.*
doi:10.1038/nature11626

Nitric oxide caught in traffic
M T Gladwin & D B Kim-Shapiro
doi:10.1038/nature11640

STRUCTURAL BIOLOGY

Structure of the Mediator head module
L Larivière *et al.*
doi:10.1038/nature11670

➲ NATURE PODCAST

This week, why redheads are doubly susceptible to cancer, the very extended bird family tree, and the genetics of the single human cell.
www.nature.com/nature/podcast
Visit our video channel to see the pesticide-threatened bees referred to on pages 43 and 105 in this issue.
www.youtube.com/NatureVideoChannel

NEWS & VIEWS

MATERIALS SCIENCE

Self-assembly gets new direction

By controlling the placement of 'sticky' patches on particles, assemblies can be made that mimic atomic bonding in molecules. This greatly expands the range of structures that can be assembled from small components. SEE ARTICLE P.51

MATTHEW R. JONES & CHAD A. MIRKIN

Imagine trying to assemble a bookcase from parts covered in glue that stick to each other equally well wherever they touch, regardless of their relative orientations. You would quickly find the task to be extremely challenging, because the components would keep joining together in haphazard configurations, rather than fitting neatly into their intended positions. Indeed, even relatively simple structures, let alone your bookcase, are impossible to create when the interactions between individual parts lack two key properties: specificity and directionality.

Scientists working with colloids — micro- and nanoscale particles suspended in a liquid — as components of self-assembling systems have found themselves in an analogous predicament. In general, the particles are spherical and uniformly sticky across their surfaces, and they interact through nonspecific forces. The lack of specificity has been addressed by attaching single-stranded DNA molecules to particles, so that they interact only with other particles bearing complementary DNA[1]. But imparting directional bonding interactions to colloidal particles has remained more of a challenge. On page 51 of this issue, Wang *et al.*[2] take the concept of DNA-mediated interactions a step further with their report of micrometre-sized particles that have symmetrically arranged, 'sticky' patches of DNA on their surfaces. The patches force the particles to interact only along certain vectors, mimicking the connectivity of atoms in molecules. This work is a major advance on earlier attempts to generate directional interactions between particles[3–12], and greatly increases the sophistication of structures that can be built 'bottom up' from smaller components.

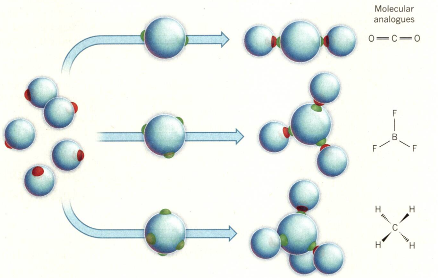


Figure 1 | Self-assembly of patchy particles. Wang *et al.*[2] report the synthesis of micrometre-sized particles that have surface patches (red and green regions) positioned to mimic the arrangements of bonds around atoms. The authors attached single-stranded DNA (not shown) to these patches, so that the resulting particles bind only to neighbouring particles whose patches bear the complementary DNA sequence; in the examples shown, red patches bind to green patches. When matching particles are mixed together, they self-assemble into clusters that resemble the atomic arrangements of molecules, such as linear carbon dioxide (CO_2), triangular boron trifluoride (BF_3) and tetrahedral methane (CH_4). In the molecular structure of methane, solid wedges indicate bonds projecting above the plane of the page, whereas broken wedges represent bonds projecting below that plane.

Directional interactions between atoms — a concept called valency — are common and form the basis for the rich structural complexity of many naturally occurring materials, from organic molecules to atomic lattices. In atoms, coordination environments (the arrangements of atoms, molecules or ions bound to a central atom) typically adopt highly symmetrical geometries, such as linear, triangular, tetrahedral or octahedral. Electron orbitals are responsible for the directionality of this bonding, which chemists regularly study and manipulate. It has long been the goal of many colloid scientists to synthesize 'artificial atoms' that interact with these same symmetries, in principle enabling man-made components to assemble predictably and with the same diversity as atoms[4,5].

Previous attempts at particle-based valency have struggled to localize sticky patches at symmetric sites and have been limited primarily to two-sided particles[6–9]. Others have used different particle shapes, such as rods, triangular prisms and octahedra, as a means to break the conventional spherical symmetry of the particle, inducing interactions that loosely mimic valency[10–12]. Wang *et al.* have vastly expanded the morphological diversity of such structures by creating particles with up to seven symmetrically positioned patches that precisely mimic atomic orbital arrangements.

To make these colloidal cousins to atoms, Wang *et al.* started with *n* polymer spheres packed into clusters with geometries that can be tailored to resemble various polyhedra[13], such as triangles ($n = 3$), tetrahedra ($n = 4$), and octahedra ($n = 6$). By swelling the clusters in a controlled way from the centre outwards — by treating them with styrene — then polymerizing the styrene, the authors made particles that had 'islands' of the original spheres protruding from the newly formed surface (see Fig. 1b, c of the paper[2]). These small, exposed regions resemble patches, and maintain the geometry of the original cluster. The authors then attached single-stranded DNA molecules to the patches, which resulted in sticky regions that mediate inter-particle binding through hybridization with complementary DNA strands attached to patches on neighbouring particles. The locations of the patches provide directionality, whereas the sequence-dependent binding of DNA imparts specificity.

With their artificial 'atoms' in hand, Wang *et al.* went on to synthesize artificial 'molecules' by combining mixtures of particles that have matched valencies and complementary DNA strands (Fig. 1). For example, when they mixed small monovalent 'B' particles (which have one patch) with a larger, tetravalent 'A' particle (which has four patches), they obtained an AB$_4$ cluster in which a central A particle was surrounded by four satellite B particles in a tetrahedral arrangement. Similarly, combinations of other appropriately matched particles yielded linear (AB$_2$) or triangular (AB$_3$) morphologies and even copolymer arrangements — long chains of alternating A and B particles. By increasing the patch size on divalent particles so that more than one monovalent particle can bind to a single patch, the authors produced even more-complex molecular motifs, such as systems that mimic the different arrangements of atoms around a double bond.

Despite their strong analogy to atoms, these patchy particles are quite large (500 to 900 nanometres in diameter), a feature that powerfully modifies the dynamics of their interactions compared with atoms. Wang and co-workers took advantage of this size difference to monitor the formation kinetics of their artificial molecules in real time using optical microscopy (for videos in the Supplementary Information to the paper, see go.nature.com/fagxzu). Such detailed kinetic studies based on direct observation are currently impossible with atoms, and so these particles might one day function as a model system to shed light on certain aspects of the dynamic and packing behaviour of matter at the smallest length scales.

In some respects, Wang and colleagues' particles are even more amenable to tailoring than naturally occurring atoms and molecules. For example, particles of different size can be combined in a multitude of configurations; the length and sequence of the DNA molecules on a given patch can modulate the spacing and connectivity between particles; and the symmetry and number of patches on a particle can be tuned to access geometries not found in any natural system. The authors' work therefore greatly expands the toolbox for assembling systems of colloidal particles. It will be surprising if some of the newly accessible configurations do not yield mechanistic insight into atomic systems, or give rise to materials that have previously unknown properties.

A major challenge for the future will be to expand Wang and co-workers' methods to generate even more sophisticated structures, for example by controlling the geometry of patches in a way that does not follow the highly symmetrical arrangements governed by simple polyhedra. Ring-like patches, linear patches and patches confined to an equatorial plane have all been shown theoretically to exhibit fascinatingly complex, self-assembled arrangements[14].

In addition, the property of chirality — the characteristic of objects that have non-superimposable mirror-image forms — may be introduced into these structures if each patch on a single particle contains a different type of DNA strand that binds to a distinct kind of particle. A tetrahedral particle could therefore capture four separate, differently sized particles, creating a cluster with a chiral centre. Control over chirality at the micro- and nanoscale is rapidly becoming a goal of great scientific interest because exotic phenomena not found in nature are thought to arise from such structural motifs[15]. Finally, extension of the principles developed by Wang *et al.* to particles of alternative compositions (such as those made from noble metals, semiconductors or oxides) will allow optical, electronic and catalytic materials to be coupled in previously impossible architectures that have potentially new emergent properties.

Wang and colleagues' work demonstrates the power of valency when applied to man-made materials, and challenges the scientific community to find methods to generalize this principle to particles at smaller and smaller length scales. With such tools in hand, scientists and engineers may one day be able to construct materials from the bottom up with a precision that makes assembling a bookcase look like child's play. ■

Matthew R. Jones *and* **Chad A. Mirkin** *are in the Departments of Materials Science and Engineering and of Chemistry, Northwestern University, Evanston, Illinois 60208, USA.*
e-mail: chadnano@northwestern.edu

1. Mirkin, C. A., Letsinger, R. L., Mucic, R. C. & Storhoff, J. J. *Nature* **382,** 607–609 (1996).
2. Wang, Y. *et al. Nature* **491,** 51–55 (2012).
3. Alivisatos, A. P. *et al. Nature* **382,** 609–611 (1996).
4. Glotzer, S. C. & Solomon, M. J. *Nature Mater.* **6,** 557–562 (2007).
5. Macfarlane, R. J. *et al. Science* **334,** 204–208 (2011).
6. Xu, X., Rosi, N. L., Wang, Y., Huo, F. & Mirkin, C. A. *J. Am. Chem. Soc.* **128,** 9286–9287 (2006).
7. Huo, F., Lytton-Jean, A. K. R. & Mirkin, C. A. *Adv. Mater.* **18,** 2304–2306 (2006).
8. DeVries, G. A. *et al. Science* **315,** 358–361 (2007).
9. Walther, A. & Muller, A. H. E. *Soft Matter* **4,** 663–668 (2008).
10. Jones, M. R. *et al. Nature Mater.* **9,** 913–917 (2010).
11. Jones, M. R., Macfarlane, R. J., Prigodich, A. E., Patel, P. C. & Mirkin, C. A. *J. Am. Chem. Soc.* **133,** 18865–18869 (2011).
12. Li, M., Schnablegger, H. & Mann, S. *Nature* **402,** 393–395 (1999).
13. Manoharan, V. N., Elsesser, M. T. & Pine, D. J. *Science* **301,** 483–487 (2003).
14. Zhang, Z. & Glotzer, S. C. *Nano Lett.* **4,** 1407–1413 (2004).
15. Soukoulis, C. M. & Wegener, M. *Nature Photon.* **5,** 523–530 (2011).

ECOLOGY

Bumblebees and pesticides

A study showing the effects of two pesticides on bumblebees highlights the need for risk assessments to consider multiple species and the complex chain of factors that determines insect exposure to chemicals. SEE LETTER P.105

JULIET L. OSBORNE

Bee and pollination science is blooming. Research efforts around the world are seeking to explain bee colony losses, which could threaten both wild and agricultural plant systems owing to the crucial role of bees in pollination. The impact of pesticides on bees is one factor that has caught the attention of scientists and the public alike: more than 100 papers and reports have been published on this topic so far this year, including research[1–3], scientific reviews[4–6] and reports on regulatory risk-assessment procedures[7–9]. On page 105 of this issue[10], Gill *et al.* strike at the heart of this debate, providing a thorough data set that examines bumblebee responses to two pesticides*.

Gill and colleagues investigated the effects of two insecticides (imidacloprid and λ-cyhalothrin) on the development and growth of bumblebee colonies, and on the foraging activity of individual bees, by tagging them with microchips. The researchers placed feeders of sugar syrup that had been spiked with imidacloprid, and/or filter paper treated with λ-cyhalothrin, in the path of bumblebees leaving their nest boxes. Significantly, the bees were not restricted to visiting the treated material — they could bypass the filter paper and the feeder, and they were able to forage in the surrounding landscape for pollen and nectar.

The authors report that fewer adult worker bees emerged from pupae in the colonies exposed to imidacloprid, which resonates with a previous study[2] that found reduced production of queen bees in imidacloprid-treated colonies. Gill and colleagues also found that bees from such colonies exhibited increased foraging activity, and that a higher proportion of foragers did not return to the colony.

---

*This article and the paper under discussion[10] were published online on 21 October 2012.

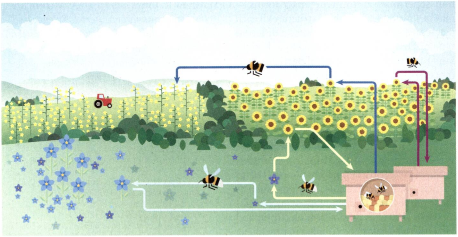


Figure 1 | A complex exposure landscape. In a typical agricultural setting, different crops may be sprayed with different pesticides at different times and doses. Bees will obtain food both from these crops and from wild plants, which makes it difficult to estimate their overall exposure to chemicals. Furthermore, bees returning to the colony after foraging may pass on the pesticides as they feed larvae. In an attempt to partially mimic this exposure complexity, Gill *et al.*[10] placed pesticide-laden feeders and filter paper (not shown) at the entrance to boxed colonies of bumblebees, which could also access flowers on crops and wild plants in the wider landscape. The researchers measured the effect of these added pesticides at both the individual-bee and colony level.

In colonies exposed to λ-cyhalothrin, they observed higher mortality of worker bees in the nest. Finally, they report that colonies exposed to both pesticides showed additive effects predictable from the individual treatment results.

This paper is important for three reasons. First, whereas most studies on bees and pesticides, and most risk-assessment models, focus on honeybees, Gill *et al.* studied bumblebees, which have a different biology and ecology and may be more vulnerable to pesticides[2]. Honeybees are smaller than bumblebees, so individual insects may be more susceptible to acute effects of chemicals, but their colonies contain tens of thousands of workers, and colony-level effects may be buffered by this sheer size. By contrast, a bumblebee colony has only a few dozen workers, so it is likely to be less resilient to the loss of individuals. The smaller colony size also makes it more difficult to monitor the survival of wild bumblebee colonies. European regulatory authorities are urgently considering how to incorporate data on bumblebees into pesticide risk assessments[3,6], and Gill and colleagues' findings provide useful input to this discussion.

Second, and unusually, the authors measured effects of pesticides on both individual bees and the whole colony. This dual evaluation stems from the concept that, if a bee acquires an acute lethal dose in the field, it will not return to the colony; however, if it ingests a sublethal dose, it might bring the material back to the colony, feed it to the brood and potentially affect the development and survival of nest-mates (an effect that would have greater impact in small colonies). Gill and colleagues' study is a helpful step towards considering this complex combination of sublethal dosage, acute and chronic mortality, and overall colony impacts, but these processes need further attention.

Examining the effects of a combination of chemicals is the third strength of this study, particularly because current regulatory requirements do not take into account the fact that insects will be exposed to multiple products[5–7]. Gill *et al.* chose chemical doses to approximate those to which a bee might be exposed in the field, although there will almost certainly be debate about whether these doses (and methods of exposure) were realistic. The chosen concentrations match label guidelines for application, but these may not reflect what farmers actually use in best or common practice.

And herein lies the catch. There are simply not sufficient field data available on the variable spatial and temporal distribution of pesticides on or in plant material, nor on bee foraging choices, to make useful comparisons between field and experimental exposure[6]. Insects probably experience a complex 'pesticide-exposure landscape' comprising multiple chemicals from several manufacturers, in one or multiple locations, applied at different doses and different times by several farmers[5] (Fig. 1). Regardless of whether the doses used by Gill and colleagues closely match field doses, their study should stimulate further exploration of the exposure landscape for bees and other non-target organisms.

The UK government and other regulatory agencies around the world are currently considering updating guidelines for pesticide registration and use. To what extent should single studies such as this one influence these decisions? Although Gill and colleagues' experimental design does not fit the normal three-tiered approach to risk assessment (which incorporates laboratory, semi-field and field experiments[7,8]), it does expose areas where the current assessment system provides insufficient evidence. Indeed, the authors' recommendations — the need for evaluations of effects on bumblebees, at individual and colony levels, and of the effects of combinations of chemicals — closely match recent recommendations from European agencies[7,8]. However, the requirement for standardization and repeatability of protocols inevitably makes for slow implementation of such recommendations. So the question remains: should policy-makers make decisions on the strength of current evidence, or should they wait for more?

Furthermore, this debate is complicated by the impact of multiple other factors on bees. For example, we have as yet no convincing demonstration of the relative effects of pesticides on bee colonies compared to the effects of parasites, pathogens and foraging resources. It is not experimentally feasible to study all possible combinations of factors in all landscapes, but modelling colony dynamics, foraging patterns and external influences is a practical and time-efficient way to make progress. Such models should be built using robust data sets and possess enough detail on life-cycle dynamics to be considered realistic[11]. The part played by mass-flowering crops, such as oilseed rape and sunflowers, also needs to be evaluated more clearly. These crops may have both positive and negative impacts on bees: they can enhance bee colony growth and pollinator-species diversity in otherwise flower-poor environments, but they are typically treated with pesticides. So the net effect on pollinator populations of growing thousands of hectares of these crops has yet tobe established.

And finally, the balance between protecting crops from pest damage and protecting pollinators needs further consideration. What alternative pest-management strategies would farmers adopt, for example, if a particular class of agrochemical were removed from their toolkit? The debate on pesticide use would be enhanced by a sound framework in which to represent the relevant socioeconomic and environmental trade-offs.

In summary, this single study does not provide a full explanation for bee declines, nor a definitive answer to questions about how to change pesticide regulations. But its convincing and detailed data set highlights the appropriateness of including bumblebees in agrochemical risk assessments and, more broadly, the need for a better understanding of pesticide-exposure landscapes. ■

Juliet L. Osborne *is at the Environment and Sustainability Institute, University of Exeter, Penryn, Cornwall TR10 9EZ, UK.*
e-mail: j.l.osborne@exeter.ac.uk

1. Henry, M. *et al. Science* **336,** 348–350 (2012).

2. Whitehorn, P. R., O'Connor, S., Wackers, F. L. & Goulson, D. *Science* **336,** 351–352 (2012).
3. Department for Environment, Food and Rural Affairs *Neonicotinoid Insecticides and Bees: The State of the Science and the Regulatory Response* (DEFRA, 2012).
4. Blacquière, T., Smagghe, G., van Gestel, C. A. M. & Mommaerts, V. *Ecotoxicology* **21,** 973–992 (2012).
5. Thompson, H. M. *Interaction Between Pesticides and Other Factors in Effects on Bees* EFSA Supporting Publs 2012:EN-340 (2012).
6. European Food Safety Authority *EFSA J.* **10,** 2752 (2012).
7. EFSA Panel on Plant Protection Products and Their Residues *EFSA J.* **10,** 2668 (2012).
8. Fischer, D. & Moriarty, T. *Pesticide Risk Assessment for Pollinators: Summary of a SETAC Pellston Workshop* Pensacola, Florida (Soc. Environ. Toxicol. Chem., 2011).
9. Oliver, R. *Am. Bee J.* **152,** 805–810 (2012).
10. Gill, R. J., Ramos-Rodriguez, O. & Raine, N. E. *Nature* **491,** 105–108 (2012).
11. Henry, M. *et al. Science* **337,** 1453 (2012).

The author declares competing financial interests.
See go.nature.com/c2a5kj for details.

ASTEROIDS

Dark and stormy weather

Can some of the ageing effects on asteroid surfaces be caused by an interplanetary rain of carbon-rich Solar System debris? Observations from the Dawn space mission suggest that the answer is yes. SEE LETTERS P.79 & P.83

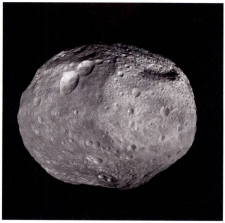
BETH ELLEN CLARK

One might expect the asteroids of our Solar System to show their age in similar ways. After all, asteroids are simply rocks that orbit the Sun through interplanetary space, and they are all subject to the same ageing processes: solar wind, micrometeorite bombardment and the occasional major impact. But asteroids do not all age in the same way, claim McCord *et al.*[1] and Pieters *et al.*[2] in this issue. Their analyses of observations of asteroid Vesta, obtained by the Dawn space mission, indicate that the asteroid's surface is not coloured with age in the same way as other bodies that, like Vesta, lack an atmosphere. Rather, Vesta shows its age by incorporating carbon-rich material from impactors (Fig. 1).

Perhaps it is not too surprising that carbonaceous material of external origin (exogenic material) is found on Vesta, one of the largest bodies in the main asteroid belt that lies between the orbits of Mars and Jupiter. Extraterrestrial spherules and micrometeorites found in Earth's stratosphere, which is just above the lowest portion of the atmosphere, have long been known to be compositionally related to carbonaceous chondrite meteorites[3]. (Chondrite meteorites contain spherules of igneous material thought to have originated in the primitive solar nebula, from which the Sun and planets formed.) However, it is surprising that the material is abundant enough to change the remotely sensed optical properties of the Vestan surface.

McCord *et al.* present compelling evidence based on Vesta's colour and brightness, as well as modelling of the putative population of impactors, to support their hypothesis that Vesta's surface is contaminated with carbonaceous material that is rich in volatile elements. Pieters *et al.* describe an analysis of the spectral differences between fresh and mature Vestan surface patches, and suggest that maturation is linked with contamination — the older the Vestan surface patch, the greater the abundance of carbonaceous exogenic material incorporated. Thus, it seems that, on Vesta, dark, carbon-rich impactor material falls to the surface and darkens the location of the impact. Due to subsequent impacts, the dark material then spreads out over time and mixes with uncontaminated surface areas.

But where does this material come from, and does it coat other asteroids in the main belt?



IDA/DLR/MPS/UCAL/JPL-CALTECH/NASA

Figure 1 | Asteroid Vesta. Analyses by McCord *et al.*[1] and Pieters *et al.*[2] indicate that the dark surface areas of asteroid Vesta, seen here in an image obtained by the Dawn space mission, are a result of the deposition of carbon-rich material from low-velocity impactors.

NATURE

50 Years Ago

The temptation to establish new scientific journals appears to be irresistible. The first number of *Radio Chimica Acta* has a foreword by a great German radiochemist, Otto Hahn, in which he poses the rhetorical question as to whether a journal, specially devoted to radiochemistry, can discharge an important function differing from that already performed by existing chemical and physical journals … He concludes that a place is needed where the results now published in a variety of journals can be brought together … To have an additional journal for papers of this kind recalls the words of the sonnet "So all is dressing old words new, spending again what is already spent". In short, interested workers must pay again to read those papers which they should reasonably be able to expect to find in the journals to which they already subscribe.
From *Nature* 3 November 1962

---

100 Years Ago

Modern Problems. By Sir Oliver Lodge — From the scientific point of view, one of the most interesting chapters is that on the smoke nuisance, in which the author deals with the problems of combustion, and advocates the use of gas fires and the suppression of crude combustion of coal in towns. As to river and sea mists, and fogs of non-avoidable kind, Sir Oliver suggests electrification of the atmosphere on a large scale … No one can tell for certain what would happen by this atmospheric electrification, but it is possible and even probable that the results might be of incalculable benefit … When we think of the tremendous harmfulness of fog … it seems obvious that the prospect of a cure of this evil would justify a large national grant for expenditure on trials in a large way.
From *Nature* 31 October 1912

The material almost certainly comes from dark asteroids in the main belt, because asteroids can be ground down to micrometeorite particles by mutual collisions over billions of years. The idea that similar exogenic material coats other large asteroids has yet to be tested. However, perplexing spectral signatures of hydrated minerals, consistent with volatile-rich carbonaceous materials, have been found on objects that would otherwise be interpreted as metallic[4,5] or inconsistent with the presence of stable volatile compounds[6]. In fact, almost a decade ago, astronomers using ground-based telescopes detected evidence of hydrated minerals on Vesta itself[7]. At the time, those measurements were considered suspect, but they now seem to be vindicated by the Dawn results.

Are the carbonaceous materials found on Vesta the result of low-velocity impactors? The fate of a rocky impactor depends in part on its velocity relative to the speed of sound in the rock. Carbonaceous materials have relatively high porosities[8]. Therefore, sound waves travel through them slowly and impact-associated shock-wave pressures may be quite low, favouring the survival of large fractions of such impactors. Without high shock-wave pressures, much less material might be vaporized and lost to space[9]. McCord *et al.* present estimates of impactor fluxes and of the amount of impactor material that Vesta has accumulated (see Supplementary Information to the paper[1]). On the basis of these estimates, they conclude that sufficient material has been deposited on the Vestan surface to cover it with a blanket up to about 1–2 metres deep.

The delivery of exogenic material is not generally what comes to mind when considering space weathering. The term is used to refer to processes that change the optical properties of the remotely sensed surface of an airless body. Studies of lunar soils and rocks brought back by the Apollo-mission astronauts have provided important information about space weathering on the Moon. Furthermore, direct evidence of space weathering on asteroids was supplied by the NEAR mission to the near-Earth asteroid 433 Eros[10] and by the Hayabusa mission to asteroid 25143 Itokawa[11].

Before the Dawn-mission findings, the consensus was that some lunar-like space weathering occurs on asteroids[12–14], and that its strength depends on the composition of the target material — that is, the material from which the asteroid is made. In the leading model of asteroidal space weathering, condensates bearing submicroscopic iron are deposited on grain surfaces after the target material has been vaporized by solar-wind sputtering and micrometeorite bombardment. Space weathering is known to cause surface darkening and spectral changes, and so these processes and their effects must be considered when interpreting the spectral properties of airless bodies. According to Pieters and colleagues, two other processes should now be considered when trying to explain the Dawn observations of Vesta: the mobility of regolith (powdery rubble that covers a planetary body) and fine-scale mixing of surface material. Therefore, the results prompt two questions. Why does Vesta not exhibit lunar-like space weathering? And was our model wrong, or do the weathering processes compete with each other?

The goal of the Dawn mission is to characterize the conditions and processes that were active during the Solar System's earliest epoch by investigating Vesta and Ceres, two of the largest asteroids that are still intact. Dawn has completed its tour of Vesta and will arrive at Ceres in February 2015 to send back data on that asteroid's low-brightness surface. It will be interesting to see whether these observations allow us to distinguish Ceres' bulk material from the 'rain' of carbonaceous material that may be contaminating its surface. ■

Beth Ellen Clark *is in the Physics Department, Ithaca College, Ithaca, New York 14850, USA, and with NASA's OSIRIS-REx asteroid sample return mission at the University of Arizona.*
e-mail: bclark@ithaca.edu

1. McCord, T. B. *et al. Nature* **491,** 83–86 (2012).
2. Pieters, C. M. *et al. Nature* **491,** 79–82 (2012).
3. Brownlee, D. E., Bates, B. & Schramm, L. *Meteorit. Planet. Sci.* **32,** 157–175 (1997).
4. Ockert-Bell, M. E. *et al. Icarus* **210,** 674–692 (2010).
5. Rivkin, A. S., Howell, E. S., Lebofsky, L. A., Clark, B. E. & Britt, D. T. *Icarus* **145,** 351–368 (2000).
6. Rivkin, A. S., Davies, J. K., Clark, B. E., Trilling, D. E. & Brown, R. H. *Lunar Planet. Sci. Conf.* **XXXII,** 1723 (2001).
7. Hasegawa, S. *et al. Geophys. Res. Lett.* **30,** 2123 (2003).
8. Britt, D. T. & Consolmagno, G. J. *Icarus* **146,** 213–219 (2000).
9. Rivkin, A. & Bottke, W. *Lunar Planet. Sci. Conf.* **XXVII,** 1077–1078 (1996).
10. Clark, B. E., Hapke, B., Pieters, C. & Britt, D. in *Asteroids III* (eds Bottke, W. F. Jr, Cellino, A., Paolicchi, P. & Binzel, R. P.) 585–599 (Univ. Arizona Press, 2002).
11. Noguchi, T. *et al. Science* **333,** 1121–1125 (2011).
12. Hapke, B. *J. Geophys. Res.* **106,** 10039–10073 (2001).
13. Pieters, C. M. *et al. Meteorit. Planet. Sci.* **35,** 1101–1107 (2000).
14. Sasaki, S., Nakamura, K., Hamabe, Y., Kurahashi, E. & Hiroi, T. *Nature* **410,** 555–557 (2001).

NEUROSCIENCE

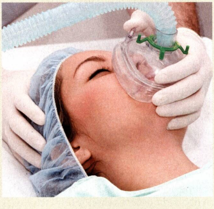
Sleep to oblivion

These days, it is hard to imagine having a surgical procedure without anaesthetics. Yet some 170 years after their first use in medicine, the way in which these drugs exert their hypnotic effects remains a mystery. Writing in *Current Biology*, Moore *et al.* shed light on the question (J. T. Moore *et al. Curr. Biol.* http://dx.doi.org/10.1016/j.cub.2012.08.042; 2012).

Many biological molecules are sensitive to anaesthetics, among them membrane ion-channel proteins. To make matters more complex, there are dozens of anaesthetic agents, and yet they don't seem to share a single molecular target. An emerging theory is that these drugs inhibit the neural circuitry associated with wakefulness. Moore and colleagues asked whether they also affect sleep-promoting neurons.

The authors focused on the anaesthetic agent isoflurane and its effects on the ventrolateral preoptic nucleus (VLPO) — a key component of the arousal (wakefulness) neurocircuitry. Neurons of the VLPO are active during sleep and, in response to inhibitory neuromodulators such as GABA, they inhibit signalling by downstream arousal-promoting neurons. Moore *et al.* find that concentrations of isoflurane that induce sedation or anaesthetic hypnosis also activate VLPO neurons, just like sleep does. Moreover, damaging these neurons reduces the hypnotic effects of isoflurane.

Of the two neuronal subpopulations that form the VLPO, only one is thought to be involved in promoting sleep. By studying slices of mouse hypothalamus — the brain region in which VLPO neurons are found — Moore and colleagues show that isoflurane specifically activates the sleep-promoting subpopulation. Exactly how it does so is unknown, but the researchers' data indicate that a reduction in the conductance of potassium ions is involved.


ISTOCKPHOTO/THINKSTOCK

The authors do not rule out a role for other sleep-promoting neuronal circuits in mediating the effects of anaesthetics. But their take-home message is that, to understand how anaesthetics act, studying sleep induction is probably just as useful as investigating the inhibition of wakefulness.

Sadaf Shadan

OPTICS

Nanotube holograms

Carbon nanotubes interact strongly with light — a property that makes them ideal components of holographic devices. The realization of such a device opens up fresh opportunities for holography.

STÉPHANE LAROUCHE & DAVID R. SMITH

Many of us grew up watching *Star Wars* and dreaming of a day when we would receive three-dimensional images as messages, like Princess Leia's urgent plea to Obi-Wan Kenobi. Although such applications are still science fiction, static three-dimensional images formed using holographic devices are real and have many practical uses, including chemical sensing, data storage and imaging. But the potential of dynamic holograms and other advanced forms continues to stimulate research into techniques and materials that might bring us closer to realizing the holograms of science fiction. Writing in *Advanced Materials*, Butt *et al.*[1] report one such development: the fabrication of a high-resolution hologram using an array of carbon nanotubes. This achievement might pave the way for holographic devices that respond to light polarization and for reconfigurable holograms.

Single-walled carbon nanotubes are hollow cylinders formed from graphene sheets — single layers of carbon atoms arranged in a hexagonal network. They can be as small as one nanometre in diameter, with lengths ranging from a few nanometres to many centimetres. Multi-walled nanotubes, such as those made by Butt *et al.*, can also be constructed, and consist of many concentric graphene cylinders.

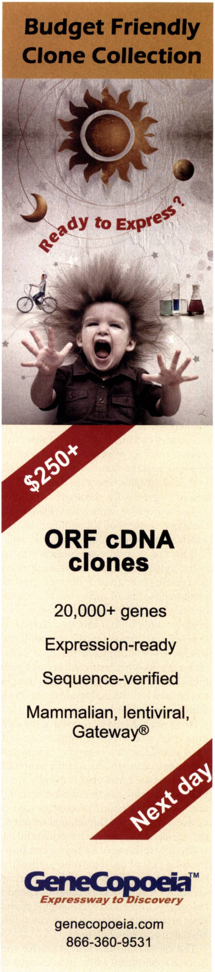
Ever since Sumio Iijima's headline-grabbing report[2] of carbon nanotubes in 1991, these tiny objects have been studied for their electrical, optical and mechanical properties. For example, depending on how the graphene sheets are wrapped, nanotubes can act as electrical conductors, semiconductors or insulators. Butt and colleagues' nanotubes behave as conductors, and therefore interact strongly with light — a characteristic that was of most interest to the authors.

Large quantities of carbon nanotubes can be produced by a technique known as arc discharge, but this yields a mixture of nanotubes, along with other carbon products, as a powder. Arc discharge is therefore ill-suited for fabricating devices (such as Butt and colleagues' holographic device) that require nanotubes of controlled dimensions located at specific positions. The authors therefore used a different synthetic approach[3,4] that allows much better control of the size and position of the nanotubes formed — a technique that had previously been used by some of the same authors to produce other optical devices[5].

First, the researchers deposited a 15-nanometre-thick layer of nickel on a silicon substrate. They then patterned the nickel using a lithographic process to create an array of dots each about 100 nm in diameter. Finally, they grew carbon nanotubes at the positions of the dots by treating the substrate with a gaseous mixture of acetylene (C_2H_2) and ammonia (NH_3) in the presence of plasma (a partially ionized gas). Under these conditions, carbon atoms from the acetylene diffuse through the nickel to form nanotubes that grow under the dots, which end up sitting on the top of the nanotubes. Meanwhile, the ammonia etches away any other carbon products that form on the rest of the substrate, leaving it pristine. This approach provided good spatial control of the resulting nanotube array: the nanotubes had an average diameter of 140 nm with a standard deviation of 13 nm, and the distance between most of them was within 25 nm of the desired spacing of 400 nm.

It is a common misconception that all holograms are three-dimensional images. On the contrary, most holograms in use are two-dimensional, or consist of a few discrete, two-dimensional layers that give some depth to an image but are not fully three-dimensional. Devices that produce two-dimensional holograms can be thought of as light-beam converters — for example, they can create a projectable image from a uniform plane wave (in which the wavefronts are parallel planes, each separated from the next by a distance of one optical wavelength). Butt and colleagues' holographic device is two-dimensional, and produces an image of the word 'CAMBRIDGE' (Fig. 1). When the device is illuminated, the incident light induces polarization — currents that oscillate in the conducting nanotubes. The polarization from each nanotube in turn generates radiation in all directions, in a process called scattering. Each nanotube can thus be thought of as a nanoscale optical antenna.

The radiation scattered from each nanotube in the device interferes constructively or destructively with the radiation scattered from all the other nanotubes depending on the direction of observation, creating bright and dark pixels in its image plane. The process of designing a hologram for the nanotube device therefore involves determining the distribution of scattering elements that is needed to produce a desired image. At distances sufficiently


Budget Friendly Clone Collection

Ready to Express?

$250+

ORF cDNA clones

20,000+ genes

Expression-ready

Sequence-verified

Mammalian, lentiviral, Gateway®

Next day

GeneCopoeia™
Expressway to Discovery

genecopoeia.com
866-360-9531

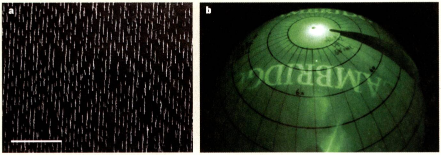


Figure 1 | Image from a nanotube array. **a**, This micrograph depicts part of Butt and colleagues' holographic device[1], which consists of a grid of carbon nanotubes. When illuminated, the presence or absence of nanotubes at each position of the grid generates a complex interference pattern that determines which pixels in the image formed by the device are light or dark; in this case, about half of the grid points contain a nanotube. Scale bar, 5 micrometres. **b**, The device generates an image of the word 'CAMBRIDGE', shown here projected on a hemispherical screen of radius 15 centimetres. (Images from ref. 1.)

far away from the holographic device, in what is called the Fraunhofer regime, the light transmission of the device and its image are simply related by a mathematical operation: the Fourier transform. This means that the arrangement of nanotubes required to produce a given image can be calculated mathematically.

To independently control the intensity of a given number of pixels in an image, one needs at least the same number of elements in the holographic device. In the case of Butt and colleagues' work, the image consisted of 300 × 300 black or white pixels. Their device therefore consists of 300 × 300 positions at which a nanotube is either present or not.

Butt *et al.* mention two reasons why carbon nanotubes are interesting components for creating holographic devices. First, the nanotubes are small and so can be patterned on a tiny grid. This is helpful because smaller grids provide holograms that have larger fields of view; a problem with currently available holographic displays is that their fields of view are somewhat restricted. Second, nanotubes interact so strongly with light that they can be used to create holograms even though their physical cross-section is small.

However, it should be noted that the nickel dots used as a template for the nanotube array would themselves produce a very similar image to that produced by the nanotubes. The role of the nanotubes is to enhance the holographic device's interaction with light and to introduce new properties into the device. For example, Butt *et al.* point out that the interaction between a nanotube and light is strongly influenced by the orientation of the long axis of the nanotube to the direction of polarization of the light. This might allow the creation of polarization-dependent holograms. One could even imagine applications in which the orientation of nanotubes is altered to activate a hologram.

The use of nanotube holograms for commercial applications will require cheaper and faster ways of producing them — Butt and colleagues' approach for creating nickel dots is well suited to research, but does not scale to mass production. However, the authors' work certainly opens up intriguing possibilities. Carbon-nanotube holographic devices made from semiconducting nanotubes would be particularly interesting, because they might enable holograms to be turned on or off electrically. Indeed, if every nanotube could be switched on and off individually, it might be possible to create reconfigurable holograms, bringing Princess Leia's message one step closer to reality. ■

Stéphane Larouche *and* **David R. Smith** *are at the Center for Metamaterials and Integrated Plasmonics, Department of Electrical and Computer Engineering, Pratt School of Engineering, Duke University, Durham, North Carolina 27708, USA.*
e-mail: stephane.larouche@duke.edu

1. Butt, H. *et al. Adv. Mater.* http://dx.doi.org/10.1002/adma.201202593 (2012).
2. Iijima, S. *Nature* **354,** 56–58 (1991).
3. Teo, K. B. K. *et al. Appl. Phys. Lett.* **79,** 1534–1536 (2001).
4. Chhowalla, M. *et al. J. Appl. Phys.* **90,** 5308–5317 (2001).
5. Butt, H., Dai, Q., Rajesekharan, R., Wilkinson, T. D. & Amaratunga, G. A. J. *ACS Nano* **5,** 9138–9143 (2011).

ORIGINS OF LIFE

The cooperative gene

The origin of life on Earth remains one of the great unsolved mysteries. A new study suggests that cooperation among molecules could have contributed to the transition from inanimate chemistry to biology. SEE ARTICLE P.72

JAMES ATTWATER & PHILIPP HOLLIGER

Cooperation operates at all scales of life, from whole organisms, such as wolves hunting in packs, to individual cells acting in a coordinated fashion during development or organ function. On page 72 of this issue, Vaidya *et al.*[1] describe networks of RNA molecules that assemble one another, suggesting that cooperation may be as old as life itself*.

The molecular architecture of modern-day organisms is based around a division of labour: the nucleic acids DNA and RNA are used mainly for the storage and processing of genetic information, with proteins fulfilling metabolic and structural roles. However, there is compelling evidence for a primordial biology that lacked DNA and proteins and instead relied on RNA for both heredity and metabolism[2]. A cornerstone of this 'RNA world' is self-replication by RNA molecules that also mutate and hence evolve towards ever more efficient self-replication.

But how did such a self-replicating RNA — the original 'selfish gene' — arise from the chemical ingredients present on the early Earth? Recent advances in prebiotic chemistry[2] (the study of the chemical reactions that might have led to the formation of the molecules typical of today's organisms) offer glimpses of how RNA's building blocks could have accumulated and polymerized into short chains[2]. Indeed, even some very short RNAs can perform chemical reactions[3] (and are therefore called RNA enzymes, or ribozymes). But it seems likely that the more complex functionalities required for self-replication would necessitate the assembly of longer, structurally more complex ribozymes, which known prebiotic reactions do not produce.

Vaidya and colleagues' remarkable work points to a possible strategy to begin bridging this gap, based on a principle of self-organization first proposed more than 30 years ago[4]. In this scenario, self-replicating RNA entities go beyond simply making copies of themselves and act on other replicators through a cyclic network of reinforcing loops called hypercycles (Fig. 1). The authors' laboratory had previously described a ribozyme — from an *Azoarcus* bacterium — that had the ability to assemble itself when fragmented[5]. Now Vaidya *et al.* show that variants of such RNA fragments can assemble and act on one another to form cooperative self-assembly cycles very much like the proposed hypercycles, in which ribozyme 1 aids assembly of ribozyme 2; 2 aids 3; and 3 aids 1 (Fig. 1).

---

*This article and the paper under discussion[1] were published online on 17 October 2012.

The authors' key finding is that, through such cooperative cycles, participating RNAs gain an advantage and can outcompete selfish replication cycles, in which a particular fragment assembles itself. Cooperation also allowed full-length ribozyme assembly from sets of four different RNA fragments. Thus, cooperation between small RNA molecules can aid the emergence of longer, more complex RNAs.

The authors describe a certain three-member cooperative cycle in great detail, but the data in one of their experiments hint at the potential for much larger cycles and networks of cooperating RNAs. This observation in particular suggests many lines of investigation that could advance our understanding of molecular cooperation and its significance to the RNA world. Questions for future enquiry include how such networks develop over time, and whether network complexity scales with efficiency — that is, whether larger or more interconnected networks always replicate more efficiently than simpler alternatives.

How might such networks have arisen (and persisted) in the pools of random RNA chains generated on the early Earth? In the present study, all members of the pool are derived from a set of 'prefabricated' fragments of the *Azoarcus* ribozyme. It will be important to determine how cooperative RNA networks perform in the presence of many unrelated and potentially interfering RNAs, and how much sequence variation within the *Azoarcus* fragments can be tolerated before self-assembly is abolished. The present work is encouraging in this respect, as it shows that limited sequence diversity in the three-member system yielded better assembly than defined fragments, demonstrating that some sequence variation can be harnessed for gains in efficiency.

Comparison with an earlier two-component system, in which two ribozymes catalysed each other's synthesis from a mixture of four fragments[6], is illustrative. This system displayed exponential self-replication and, when seeded with fragments of defined sequence variation, yielded a diverse pool of recombinant molecules, some of which were more efficient replicators than the initial ones. Thus, such molecular systems can harness the powerful evolutionary potential of recombination to reassort themselves into more active replicators. Vaidya and colleagues' use of a ribozyme system that had a larger degree of freedom in the choice of assembly partners has now enabled networks to develop beyond this two-component system.

However, the need for defined RNA components is likely to constrain the evolutionary potential of such systems, because recombinants are unable to break away from the prescribed component structure. A more general capacity for self-replication and evolution would require a different type of system that has the ability to copy genetic information — akin to present-day biology, in which RNA or DNA sequences are replicated by polymerase enzymes 'letter by letter' from monomer units. Although RNA-polymerizing ribozymes have been described[7], their activity falls short of self-replication, despite recent improvements[8].

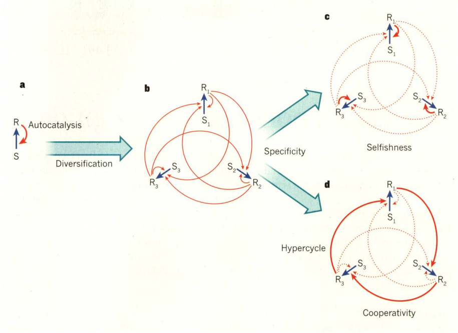


Figure 1 | The emergence of hypercycles. a, A primordial replicator molecule (R) enhances its own assembly from substrate molecules (S) in a simple autocatalytic cycle. **b**, Imperfect replication generates a set of related replicators, each promoting the synthesis of all the others. **c**, **d**, The introduction of biases in replicator specificity gives structure to the network and can lead to selfish subsystems (**c**) or to a cooperative 'hypercycle' (**d**), akin to the system described by Vaidya and colleagues[1]. Such hypercycles remain globally autocatalytic, but are more resistant to the accumulation of mutations, enabling replicators to specialize and to acquire new functions. Thick and dashed red arrows indicate increased and decreased efficacy, respectively, at enhancing replicator assembly.

The excursions into 'molecular ecology' described by Vaidya *et al.* suggest that cooperative networks might be designed to harness the best of both types of system, if synthesis of short RNAs by polymerizing ribozymes could be coupled to a ribozyme system capable of self-assembly[9]. Such networks might outperform replicators that go it alone, and exploit recombination to resist the gradual accumulation of harmful mutations[10] and the concomitant deterioration of the encoded genetic information. Finally, complete covalent assembly might not be essential for higher-order functions such as molecular kin recognition and polymerizing activities. Indeed, non-covalent assembly of multiple RNA chains into functional complexes has precedents in modern-day biology, notably the ribosome, a large complex of multiple RNA and protein chains, which catalyses protein synthesis and may date back to the RNA world.

The precise molecular events that led to the origin of life on Earth are likely to be lost in time, but science can construct molecular 'doppelgängers' of the ancestral molecules and explore the plausibility of different ways in which the transition from prebiotic to biotic matter might have occurred. Vaidya and colleagues make a persuasive case for the benefits of cooperation even at this nascent stage of life. The first genes may not have been so selfish, after all. ■

James Attwater *and* **Philipp Holliger** *are at the MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK.*
e-mail: ph1@mrc-lmb.cam.ac.uk

1. Vaidya, N. *et al. Nature* **491,** 72–77 (2012).
2. Atkins, J. F., Gesteland, R. F. & Cech, T. R. (eds) *RNA Worlds: From Life's Origins to Diversity in Gene Regulation* (Cold Spring Harb. Lab. Press, 2011).
3. Turk, R. M., Chumachenko, N. V. & Yarus, M. *Proc. Natl Acad. Sci. USA* **107,** 4585–4589 (2010).
4. Eigen, M. & Schuster, P. *Naturwissenschaften* **65,** 341–369 (1978).
5. Hayden, E. J. & Lehman, N. *Chem. Biol.* **13,** 909–918 (2006).
6. Lincoln, T. A. & Joyce, G. F. *Science* **323,** 1229–1232 (2009).
7. Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E. & Bartel, D. P. *Science* **292,** 1319–1325 (2001).
8. Wochner, A., Attwater, J., Coulson, A. & Holliger, P. *Science* **332,** 209–212 (2011).
9. Meyer, A. J., Ellefson, J. W. & Ellington, A. D. *Acc. Chem. Res.* http://dx.doi.org/10.1021/ar200325v (2012).
10. Muller, H. J. *Genetics* **48,** 903 (1963).

nature

1 November 2012 / Vol 491 / Issue No. 7422

Publication policies, author guidelines and information

This is a summary of *Nature*'s guide to authors, publication policies and other author information, available in full on the web via our main author information page at www.nature.com/nature/authors.

Getting published in *Nature*

Please read www.nature.com/nature/authors/get_published before finalizing any original scientific research manuscripts you intend to submit. This section explains *Nature*'s editorial criteria, and how manuscripts are handled by the editors after submission and before acceptance for publication.

Formatting guide

Please read www.nature.com/nature/authors/gta before finalizing and submitting your contribution. This section provides instructions for preparing, formatting and writing all types of contribution published by *Nature*.

Publication policies

This section, at www.nature.com/authors, describes *Nature*'s conditions of publication, including policies on authorship; competing financial interests; duplicate publication; prepublicity and embargoes; materials and data availability; research on animal (including human) subjects; referees; corrections, retractions, errata and addenda to published papers; the complaints procedure and refutations; and copyright/licence to publish. All contributions, for whatever section of the journal, are considered for publication and published on the condition that authors agree to these policies.

Submissions

These sections, at www.nature.com/nature/authors/submissions, contain our online service for presubmission enquiries and submission and tracking of Articles, Letters and Review-type manuscripts. Authors should read the formatting guide before using this service. Authors who cannot submit online can submit by post, but *Nature* strongly recommends the online service. **Authors must not submit by both methods.** Authors experiencing operational difficulty in using the online submission service can send an enquiry to mtshelp@nature.com, with a brief explanation of the problem.

Editorial contacts and author services

This section, at www.nature.com/nature/about/contact, provides the addresses of *Nature*'s editorial offices; names of editorial staff; and contact details for advice, reprints, permissions, press office (press@nature.com), and other author-related contact information, including postal addresses for mailed submissions. Short biographies of editors are also available at www.nature.com/nature/about/editors.

Final submissions

This section, at www.nature.com/nature/authors/submissions/final, applies only if an Article or Letter has been accepted in principle for publication and the author has been asked by the editor for a final version. Authors will need to upload electronic files of the text, production-quality figures and supplementary information (if any) to our database. Authors must include a completed manuscript checklist and all relevant signed forms and declarations as requested by the editor.

Short version of guide to authors

See www.nature.com/nature/authors for full version. Contributors not used to *Nature*'s format are advised to read some papers in published issues before submission, to familiarize themselves with *Nature*'s style. *Nature*'s main formats for original research are Articles and Letters.

ARTICLES are original reports whose conclusions represent a substantial advance in understanding of an important problem and have immediate, far-reaching implications. They do not normally exceed **5 pages** of *Nature* and have **no more than 50** references. (One page of undiluted text is about 1,300 words.)

Articles have a summary for readers outside the discipline, separate from the main text and containing no references, of up to 150 words. The summary does not contain numbers, abbreviations, acronyms or measurements unless essential. It contains a brief account of the background and rationale of the work, followed by a statement of the main conclusions, introduced by the phrase 'Here we show' or its equivalent.

The Article itself typically has about 3,000 words of text, beginning with up to 500 words of referenced text expanding on the background to the work, followed by a concise, focused account of the findings, and ending with one or two short paragraphs of discussion. The text can contain a few short subheadings consisting of about 20 but no more than 40 characters (including spaces) each.

LETTERS are short reports of original research focused on an outstanding finding whose importance means that it will be of interest to scientists in other fields. They do not normally exceed **4 pages** of *Nature*, and have **no more than 30** references. They begin with a **fully referenced paragraph** of not more than 300 words, aimed at readers in other disciplines. This paragraph contains a summary of the background and rationale for the work, followed by a one-sentence statement of the main conclusions introduced by the phrase 'Here we show' or its equivalent, and ending with two to three sentences describing the implications in context.

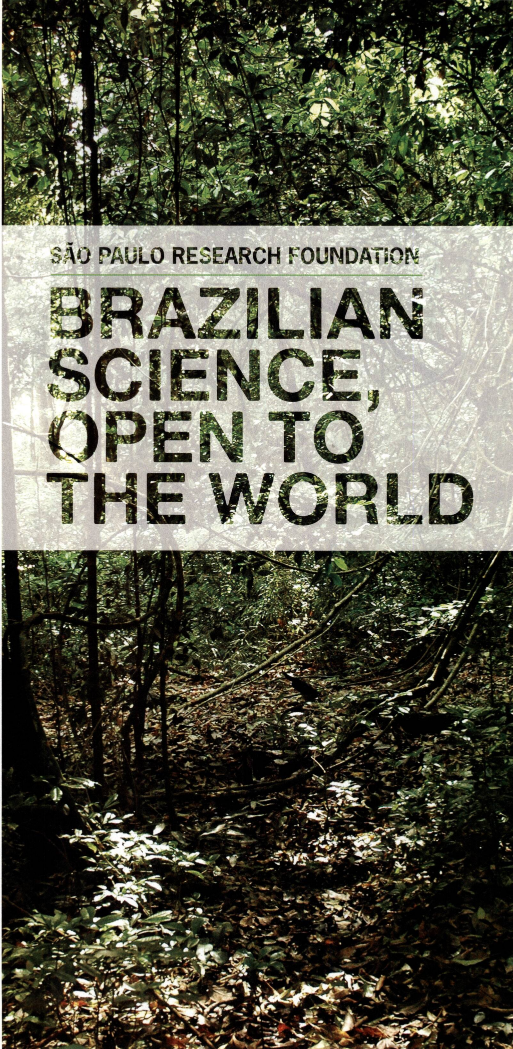
The main text which follows is typically up to 1,500 words long. This text can, if the author requires, start with a further brief paragraph of introductory material. A description of the research then follows. Any discussion at the end of the text should be as succinct as possible.

FIGURES in both Articles and Letters are as small as possible: Articles typically have 5 or 6 figures, Letters typically have 3 or 4 figures. Legends and/or methods are additional to the main text. Methods Summaries (which are published in print only) are no longer than 300 words; Methods sections (which are published online) can be up to 1,000 words; legends should not exceed 100 words each and start with a one-sentence title for the figure, followed by a brief explanation of the parts and symbols.

See www.nature.com/nature/authors/gta/others.html for details of other types of contribution to *Nature*, including Brief Communications Arising; Reviews and Perspectives (presubmission enquiries to reviewsandinsights@nature.com); and Correspondence (submit to corres@nature.com).

See Authors and Referees @ NPG (www.nature.com/authors) for details of Nature journals' publication policies, author and referee services, author benefits and for links to Nautilus and Peer to Peer blogs.

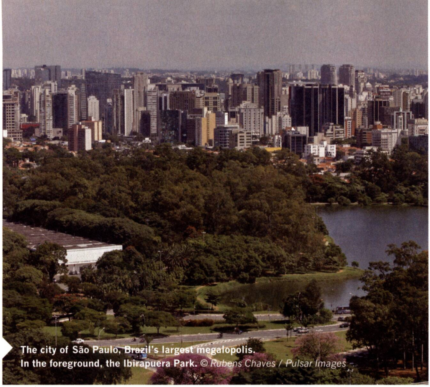
ADVERTISING FEATURE



SÃO PAULO RESEARCH FOUNDATION

BRAZILIAN SCIENCE, OPEN TO THE WORLD

FAPESP



The city of São Paulo, Brazil's largest megalopolis. In the foreground, the Ibirapuera Park. © *Rubens Chaves / Pulsar Images*

HALF OF BRAZIL'S SCIENCE OUTPUT IS FROM SÃO PAULO

For 50 years, FAPESP has been supporting research in the State

Brazil is emerging as a major contributor to scientific knowledge. In 2011, the country stood fifteenth in terms of the number of scientific papers published, a position that will be boosted in years to come by the twelve thousand new doctorates that are graduating annually.

More than half of the country's scientific output is concentrated in the State of São Paulo, which is also Brazil's economic powerhouse. With almost 42 million inhabitants, the state accounts for 35 percent of Brazil's gross domestic product (GDP), outstripping countries like Belgium, Norway, Poland, and even the Netherlands.

At the heart of the state's progress in science is FAPESP, the São Paulo Research Foundation. Created in 1962, FAPESP is a dynamic research-funding agency. Over the past half century, it has supported the training of 112,000 scientists from undergraduate to postdoctoral level, provided financial backing to 96,000 individual and thematic research projects, and funded substantial improvements to the research infrastructure.

Currently, more scientific articles are published by researchers from the state of São Paulo than from any country in Latin America, other than Brazil as a whole. And, with 1.52 percent of GDP being invested in R&D, a figure that surpasses Spain, Italy, Portugal, Mexico, Argentina and Chile, the science sector in the state will continue to develop strongly.

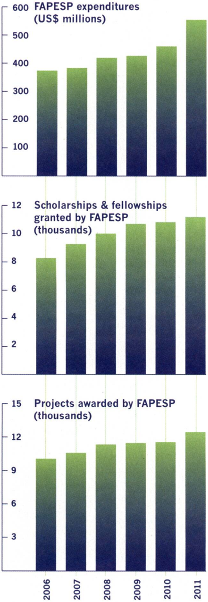
Five major state- and federally-funded universities are located in the State of São Paulo, The University of São Paulo (USP), The State University of Campinas (UNICAMP), The São Paulo State University (UNESP), The Federal University of São Paulo (UNIFESP), The Federal University of ABC (UFABC) and The Federal University of São Carlos (UFSCar).

The three state universities – USP, UNICAMP and UNESP – have particularly strong graduate programs. Every year, two thousand doctorates graduate from USP, eight hundred from UNICAMP and six hundred from UNESP.

The state also hosts nineteen research institutes, including the Aeronautics Technology Institute (ITA), the National Space Research Institute (INPE) and the National Synchrotron Light Laboratory (LNLS)

US$554 MILLION FOR SCIENCE IN 2011

FAPESP enjoys full administrative and financial control over one percent of the State's tax revenue, guaranteed by the State Constitution. In 2011, this generated a research budget for FAPESP of some US$554 million to support all areas of fundamental and applied science as well as technology, engineering, humanities and arts. Evaluations for funding are based on rigorous and confidential peer review of scientific merit.



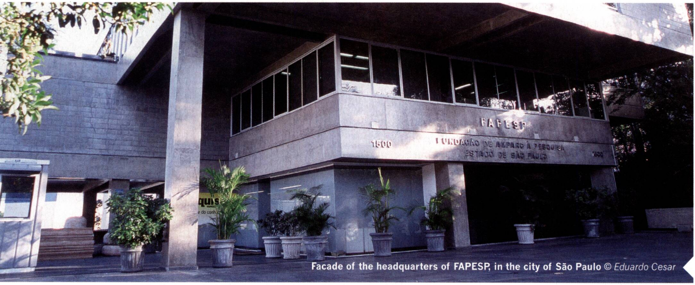
FAPESP blossomed in the 1990's, after the end of military rule (1964-1985) and the restoration of democracy in Brazil, and reached full maturity with a highly successful genome program. This began in 1997 with the sequencing of the bacterium *Xylella fastidiosa* (which causes the disease of citrus fruits known as "yellow"), a study that made the cover of *Nature*. It continued with the sugarcane genome program, which was initiated in 1998 and positioned Brazilian science at the forefront of genomics.

Almost 250 researchers around the country participated in the Sugarcane Genome Project, identifying nearly 50,000 genes involved in growth, sugar production, disease resistance and survival in adverse climate and soil conditions. The sequencing of the *Xanthomonas citri* bacterium, which causes citrus canker, another disease of orange trees, and the Cancer Genome Project, were also great successes of the Genome Program.

With the FAPESP Genome Program, the Foundation proved itself capable of organizing large, multidisciplinary research programs that have immediate economic and social impact. It provided first-hand experience of the high degree of integration required for scientific and technological research on a global scale and introduced the country's research base to international cooperation. Building on this success, FAPESP is committed to establishing partnerships and agreements with funding agencies, higher education and research organizations, and companies from different countries.

The Foundation seeks to promote research in areas considered strategic for the country and crucial for the advance of world science. Ambitious programs are being supported in bioenergy (BIOEN), biodiversity (BIOTA) and global climate change (RPGCC), among other areas. » **www.fapesp.br/en**

ON THE COVER: The city of São Paulo, capital of the state of São Paulo. © *Delfim Martins / Pulsar Images* **Atlantic Forest, protected reservation in the municipality of Gália, state of São Paulo.** © *Eduardo Cesar*



Facade of the headquarters of FAPESP, in the city of São Paulo © *Eduardo Cesar*

INTERNATIONAL PARTNERSHIPS & AGREEMENTS

UNITED STATES: National Science Foundation; Fulbright Foundation – Dr. Ruth Cardoso Program; Oak Ridge National Laboratory; Massachusetts Institute of Technology; North Carolina State University; University of California – Dr. Ruth Cardoso Program; University of Southern California; Agilent Technologies; Boeing; Microsoft Research.

CANADA: Agence Universitaire de la Francophonie; International Science and Technology Partnerships; National Sciences and Engineering Research Council of Canada; Consortium of Alberta, Laval, Dalhousie and Ottawa; McMaster University; University of Ontario Institute of Technology; Simon Fraser, Concordia, York and Ryerson Universities; Toronto and Western Ontario Universities; University of Victoria.

UNITED KINGDOM: British Council; Research Councils United Kingdom; King´s College London; University of Nottingham; University of Surrey; University of Southampton; Bangor University; Institute of Education University of London; University of Edinburgh; University of York; Imprimatur Capital.

FRANCE: Agence Nationale de la Recherche; Centre National de la Recherche Scientifique; Région Provence-Alpes-Côte d´Azur; Centre de Coopération Internationale en Recherche Agronomique pour le Développement; Ecole Normale Supérieure; Institut National de la Recherche Agronomique; Institut National de Recherche em Informatique et en Automatique; Institut National de la Santé et de la Recherche Médicale; Paris-Tech.

EUROPE: Worldwide LHC Computing Grid.

GERMANY: DAAD, German Academic Exchange Service; DFG, Deutsche Forschungsgemeinschaft; and STMWFK, The State Ministry of Sciences, Research and The Arts of the Free State of Bavaria .

NETHERLANDS: Bio-based Ecologically Balanced Sustainable Industrial Chemistry; Erasmus Universiteit Rotteerdam.

DENMARK (Danish Council for Strategic Research).

SPAIN (Universidad de Salamanca).

ARGENTINA (Consejo Nacional de Investigaciones Científicas y Técnicas).

MEXICO: Instituto de Innovación de Nuevo León

ISRAEL (The Hebrew University of Jerusalem; Tel Aviv University).

» **www.fapesp.br/en/agreements**

ADVERTISER RETAINS SOLE RESPONSIBILITY FOR CONTENT

»FAPESP BIOENERGY RESEARCH PROGRAM (BIOEN)

Clean, renewable resources already provide 46 percent of the total domestic energy matrix. Today's research program will yield further gains.

Brazil is internationally recognized for leadership in the use of bioethanol as an alternative to gasoline. Today, sugarcane ethanol provides more than 40 percent of the fuel consumed by light vehicles in the country and the thermal energy and electricity produced using ethanol and sugarcane bagasse, which is the fibrous matter that remains after stalks are crushed to extract their juice, meets 18 percent of domestic energy needs. Due to these and other factors, the Brazilian energy grid provides an exemplary case of sustainability: 46 percent of the country's energy matrix is generated from clean, renewable resources, compared with an average of just 13 percent for all countries.

Favorable soil and climate conditions mean that Brazil is the largest global producer of ethanol and sugar from sugarcane. Currently, sugarcane is cultivated on 7 million hectares (ha), which is 2 to 3 percent of the total area devoted to agriculture. The country is also the world leader in production technology as a result of long-term R&D programs in universities, research institutions and the private sector, where both growers and mill owners have supported research. At present, the average sugarcane yield at is 75 tons/ha but recent studies indicate a theoretical potential of 380 tons/ha. Biotechnology research and the development of more efficient agricultural practices will be applied to drive gains in yield.

The program team numbers over 300 researchers from 16 institutions of higher education and research.

The operating model for the industry is to produce sugar for bioethanol production and to generate bioelectricity, which can be sold to the national energy grid. Modern production plants are increasing output and lowering costs by using the bagasse surplus to increase ethanol production without increasing the cultivated area.

While the Brazilian sugar and ethanol industry leads the world, great challenges lie ahead in the future. Second and third generation biofuels, as well as environmental and sustainability issues, demand continuous research and the development of novel tools. FAPESP, through its BIOEN Research Program, is an important part of the Brazil's endeavor to maintain leadership in clean and sustainable bioenergy.

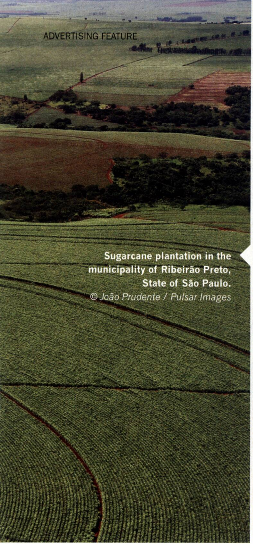
INTEGRATING PUBLIC AND PRIVATE R&D

Begun in 2008, the FAPESP Bioenergy Research Program (BIOEN) stimulates and coordinates research and development in academic and industrial settings to advance knowledge and its application in areas related to bioenergy production. The Program is organized into five divisions:

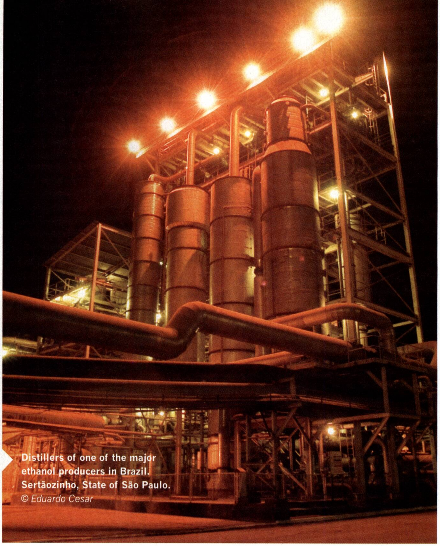
- Biomass Research (including plant improvement and sugarcane farming.)
- Ethanol and Biofuel Industrial Technologies.
- Biorefinery Technologies, Alcohol Chemistry, Sugar Chemistry and Oil Chemistry.
- Ethanol Applications for Motor Vehicles and Aviation Fuels.
- Research on Impacts, specifically social, economic and environmental studies; land use; intellectual property, institutional bottlenecks of a bio-based economy.

The program team numbers over 300 researchers from 16 institutions of higher education and research in the state of São Paulo. It is guided by six strategic goals, which are:

- To increase sugarcane yield through the use of novel practices, breeding and biotechnology.
- To increase ethanol production through optimization of existing, first generation process and the development and implementation of new technologies for first, second and third generation ethanol production processes.
- To increase the production of bio-based chemicals for the substitution of petrochemicals. To evaluate and mitigate



Sugarcane plantation in the municipality of Ribeirão Preto, State of São Paulo.
© João Prudente / Pulsar Images



Distillers of one of the major ethanol producers in Brazil. Sertãozinho, State of São Paulo.
© Eduardo Cesar

the environmental, social and economic impact of bioenergy production.

- To develop a suitable platform for biorefineries based on sugarcane and ethanol as feedstock.
- To generate knowledge that guarantees Brazil's leading position in the global bioenergy research and industry.
- In partnership with private companies, BIOEN's research agenda includes biomass production and processing; the production of biofuels; the development of engines; alcohol chemistry applications; and environmental and socioeconomic impacts. The program supports 84 basic research projects and applications for sustainable production of biofuels that are based on, but not limited to, sugarcane ethanol. Fifteen of the projects include research collaborations with other countries. » **www.fapesp.br/en/bioen**
» **www.bioenfapesp.org**

» BIOTA-FAPESP PROGRAM

1,800 new species of plants and animals have been catalogued.

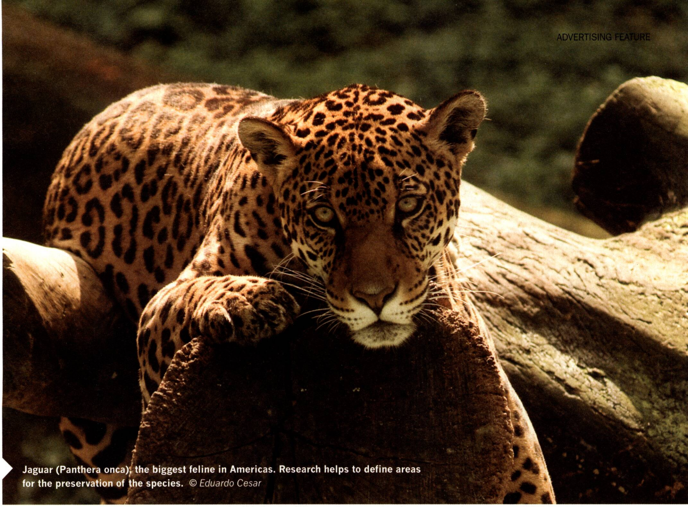
Brazil has incredible natural resources. Home to between 15 and 20 percent of the planet's biodiversity, the country encompasses six large biomes — Amazonia, Caatinga (scrubland), Atlantic Forest, Cerrado (savannah), Pantanal (wetlands) and Pampa (grassland) — originally distributed across 8.5 million square kilometers of the continent and more than eight thousand kilometers of coastline. Despite intense urbanization and the high density of occupation and use of the land, the State of São Paulo, located in southeast region of Brazil, possesses an extremely rich biological diversity. For example, the number of plant species within São Paulo territory amounts to two-thirds of that found in the entire European continent.

The BIOTA-FAPESP Program was created in 1999 with the principal objective of cataloguing and characterizing the biodiversity of the State of São Paulo, defining the mechanisms for its conservation, evaluating its economic potential and stimulating its sustainable use.

For more than 10 years, the BIOTA-FAPESP Program has been a major conservation initiative that is founded on a solid scientific basis. Research projects include:

- The molecular genetics for classification of species.
- Evolution studies to understand the origin of processes that generate, conserve or reduce biodiversity.
- Investigation into human dimensions of conservation and sustainable usage.

ADVERTISER RETAINS SOLE RESPONSIBILITY FOR CONTENT



Jaguar (Panthera onca), the biggest feline in Americas. Research helps to define areas for the preservation of the species. *© Eduardo Cesar*

Based on knowledge generated by the BIOTA program, the São Paulo legislation has made 14 decisions on conserving threatened biomes.

Scientists participating in the program have published nearly 900 scientific articles, 20 books and two atlases, and have catalogued more than 1,800 new species of plants and animals. The program has drafted maps which identify conservation areas based on 151,000 records of more than 9,000 species. Based on this body of knowledge, the São Paulo legislation has made 14 decisions on conserving threatened biomes.

The databases of the Biota-FAPESP Program are open to the scientific community. The Environmental Information System, SinBiota (www.sinbiota.biota.org.br), registers and integrates the geographical coordinates of thousands of species of plants and animals in the State of São Paulo. This database can be queried using the scientific name of a plant or animal, the name of the collector, or the locality or date of collection.

It is also possible to see the geographical distribution of species since SinBiota is a cartographic-based system that details remnants of native vegetation, areas reforested with exotic species (such as Pinnus and Eucalyptus), conservation units, the river and road networks, and the urban areas. A second database, the Biota-FAPESP Program Atlas incorporates the Forest Inventory of São Paulo, which is a survey coordinated by the Forestry Institute. Generated using field surveys, aerial photos and satellite images, the inventory monitors the area occupied by remnants of native vegetation in the State of São Paulo.

A third database, Species Link (www.splink.cria.org.br), has aggregated two million records from research activities or imported from the archives of national and foreign biological collections.

Other developments of Biota-FAPESP are the electronic scientific magazine Biota Neotropica (www.biotaneotropica.org.br), which publishes studies on biodiversity in the Neotropical region, and the Biota Network of Bioprospecting and Biotrials (www.bioprospecta.org.br), which integrates the work of research groups in the State of São Paulo that are engaged in economic development projects based on microorganisms, macroscopic fungi, plants, invertebrates (including marine) and vertebrates.

In 2009, FAPESP renewed funding of the BIOTA-FAPESP program for another 10 years, with a view to prolonging and enhancing the results and scope of the work. Named "BIOTA + 10", one of the strategic goals of the program is the expansion of multidisciplinary research on marine biodiversity. Brazil has a coastline of around 8,000 km and an adjacent continental shelf of over 800,000 km^2. This contains an array of ecosystems and environmental settings that harbor substantial marine biodiversity, nowadays referred as "Blue Amazonia." Currently, little is known about this fantastic biological resource.

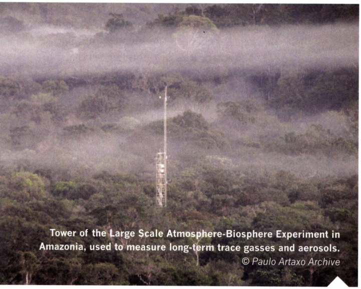
Among the topics proposed for BIOTA + 10 are inventories of marine organisms and the identification of environmental factors that affect their distribution and abundance; studies of marine ecosystems and the impact of oil and gas exploitation upon them; and analyses of marine biodiversity that might be incorporated into climatic change models.

One goal of such a marine inventory is the provision of data to inform laws regarding the protection of diverse marine biomes. At the same time, the program also emphasizes bioprospecting for marine natural products – mainly bioactive compounds for pharmaceutical production. The research will help to promote informed debate on the sustainable use of Brazilian marine biodiversity resources.

» **www.fapesp.br/en/biota**

» FAPESP RESEARCH PROGRAM ON GLOBAL CLIMATE CHANGE (RPGCC)

The goal is to develop a Brazilian model of the global climate system, by 2013.



Tower of the Large Scale Atmosphere-Biosphere Experiment in Amazonia, used to measure long-term trace gasses and aerosols.
© Paulo Artaxo Archive

There are four key drivers of climate change research in Brazil. First is the continental-sized dimensions of the country. Second is the dependence of the economy on renewable natural resources. Third is the fact that the vegetal cover has been significantly altered in the Atlantic Forest and the Cerrado, and is quickly being replaced in Amazonia. And fourth is the increased emissions of greenhouse gases and aerosols resulting from the combustion of fossil fuels and industrial and agricultural processes. A bold study that encompasses and relates all of these variables is essential for public policy in areas such as risk assessment, mitigation of effects, and adaptation to the changes.

The FAPESP Research Program on Global Climate Change (RPGCC) is just such an initiative, supporting multidisciplinary research projects for up to six years. One of the major goals is a Brazilian model of the global climatic system by 2013, with a focus on key regions including the Amazon, the Cerrado and the South Atlantic regions.

The Program combines observation and modeling components, and includes long-term environmental measurements as well as the generation of paleoclimate data, a crucial requirement to overcome the lack of long-term environmental observations. The FAPESP program is delivered in association with other funding mechanisms, both internal and external to the State of São Paulo.

The RPGCC also has a substantial technological component for the development of the appropriate technologies for a sustainable future. This encompasses innovative technologies for the mitigation of emissions and adaptation in all sectors and economic activities, given that climate change is inevitable and societies must adapt their socio-economical systems to these changes. The Program also includes a research component on the interface between earth science and climate policy.

The RPGCC considers proposals for research which fall within the scope of the Program described above, in particular in the following areas:

- Consequences of global climate change to the functioning of ecosystems, with emphasis on biodiversity loss and water, carbon and nitrogen cycles.
- Changes in the atmospheric radiation balance, including the effects of aerosol particles, clouds, trace gases and land use change.
- The impact of global climate change on agriculture productivity and food security.
- Energy production and the mitigation of greenhouse gases emissions.
- The impact of climate change on human health.
- Human dimensions of global climate changes: impacts, vulnerabilities and social and economic responses, including adaptation strategies.

Development of the Program is facilitated by a high performance supercomputer, provided jointly by FAPESP and the Science and Technology Ministry. The massive processing capacity positions the program among the world's most important for climate studies, with the capacity to create and analyze global climate models. The equipment—named Tupã, in reference to the god of thunder in one of the native Brazilian mythologies—has a peak performance of 244 teraflops

ADVERTISER RETAINS SOLE RESPONSIBILITY FOR CONTENT

(trillions of floating point operations per second).* The system incorporates recent advances in numerical modeling, climate change modeling, data assimilation, chemicals and aerosols, atmosphere, oceans and vegetation.

The model will be used to analyze interactions between the multiple elements of the terrestrial system to understand the influence of anthropogenic actions – such as greenhouse gas and aerosol emissions, changes in vegetation and the effects of urbanization on climate. Studies on the nitrogen and carbon cycles in nature and their impact on agriculture and livestock are also among the research topics.

The environment of the Amazon is an extremely important aspect of the project. It represents a very dynamic and complex system, and, to understand it, researchers need to comprehend the main processes that govern its functioning, such as the carbon balance and changes in the aquatic systems. » **www.fapesp.br/en/rpgcc**

**The flop (floating point operations per second) is the unit of computer performance. For comparison, an ordinary four operations calculator has a performance of 10 flops.*

» OPPORTUNITIES FOR POSTDOCS & YOUNG RESEARCHERS

Young researchers are strongly encouraged by FAPESP programs. © *Eduardo Cesar*

FAPESP has ongoing programs and support mechanisms to attract scientific researchers from overseas. The research positions are in a wide range of projects developed in the São Paulo Centers of Excellence. The greatest number of opportunities for foreign scholars to participate in postdoctoral research is in Thematic Projects in which experienced scientists, many of whom are renowned leaders in their field, collaborate for up to six years. In 2011, FAPESP spent US$ 45 million supporting Thematic Projects. Scholarships of up to US$ 34,000 annually for up to four years are available in all Thematic Project areas. A list of the 1,452 Thematic Projects that have been or are currently supported by the foundation can be found on the website: » **www.bv.fapesp.br/en/1/thematic-projects.**

Information regarding open opportunities can be accessed at: » **www.fapesp.br/oportunidades.**

Young Investigators Awards also offer funding of up to US$ 38,000 per annum. Recent graduates who have produced relevant scientific findings and who have feasible, solid and innovative proposals can use the award to collaborate with research teams at existing São Paulo centers or can form new groups in institutions in the State that do not have a research tradition in the particular subject. In 2011, FAPESP's commited US$ 21 million to the program. Further information is available at: » **www.fapesp.br/en/yia.**

ADVERTISER RETAINS SOLE RESPONSIBILITY FOR CONTENT

doi:10.1038/nature11564

Colloids with valence and specific directional bonding

Yufeng Wang[1], Yu Wang[1], Dana R. Breed[2], Vinothan N. Manoharan[3,4], Lang Feng[5], Andrew D. Hollingsworth[5], Marcus Weck[1] & David J. Pine[5]

The ability to design and assemble three-dimensional structures from colloidal particles is limited by the absence of specific directional bonds. As a result, complex or low-coordination structures, common in atomic and molecular systems, are rare in the colloidal domain. Here we demonstrate a general method for creating the colloidal analogues of atoms with valence: colloidal particles with chemically distinct surface patches that imitate hybridized atomic orbitals, including *sp*, sp^2, sp^3, sp^3d, sp^3d^2 and sp^3d^3. Functionalized with DNA with single-stranded sticky ends, patches on different particles can form highly directional bonds through programmable, specific and reversible DNA hybridization. These features allow the particles to self-assemble into 'colloidal molecules' with triangular, tetrahedral and other bonding symmetries, and should also give access to a rich variety of new microstructured colloidal materials.

The past decade has seen an explosion in the kinds of colloidal particles that can be synthesized[1,2], with many new shapes, such as cubes[3], clusters of spheres[4–6] and dimpled particles[7,8] reported. Because the self-assembly of these particles is largely controlled by their geometry, only a few relatively simple crystals have been made: face-centred and body-centred cubic crystals and variants[9]. Colloidal alloys increase the diversity of structures[10–12], but many structures remain difficult or impossible to make. For example, the diamond lattice, predicted more than 20 years ago to have a full three-dimensional photonic bandgap[13], still cannot be made by colloidal self-assembly because it requires fourfold coordination. Without directional bonds, such low-coordination states are unstable.

Unlike colloids, atoms and molecules control their assembly and packing through valence. In molecules such as methane (CH_4), the valence orbitals of the carbon atom adopt sp^3 hybridization and form four equivalent C–H bonds in a tetrahedral arrangement. In the colloidal domain, the kinds of structures that could be made would vastly increase if particles with controlled symmetries and highly directional interactions were available. What is needed are colloids with valence[14].

One approach is to decorate the surface of colloidal particles with 'sticky patches' made of synthetic organic or biological molecules (for example) and assigned to specific locations[15–19]. Bonding between particles occurs through patch–patch interactions, so that in principle the location and functionality of the patches can endow particles with bonding directionality and valence. This approach is conceptually simple, yet challenging to realize. For example, so-called Janus particles with asymmetrically functionalized surfaces can be made, but normally have only a single patch[20–22]. Triblock Janus particles have also been fabricated by glancing-angle deposition and assembled into a kagome lattice, the two-dimensional analogue of a diamond crystal[23]. However, only two patches are made using this method, and low quantities of particles are obtained. Other strategies have used faceted particles[24], particles with protrusions[25] or coordinated patches[26–28], but three-dimensional directional bonding and assembly have yet to be demonstrated[29].

Here we demonstrate the synthesis and assembly of colloidal particles with directional interactions that mimic those of atoms with either monovalent *s* or *p* orbitals, or multivalent *sp*, sp^2, sp^3, sp^3d, sp^3d^2 or sp^3d^3 hybridized orbitals. We do so by making particles with various numbers of patches, $n = 1$–7 and higher, that adopt spherical, linear, triangular, tetrahedral, trigonal dipyramidal, octahedral or pentagonal dipyramidal symmetries. The patches are then site-specifically coated with oligonucleotides, enabling a reversible and controllable attraction between patches on different particles. Using these colloidal 'atoms', we demonstrate that a vast collection of colloidal molecules and macromolecules are readily accessible through self-assembly schemes that are analogous to chemical reactions.

Synthesis

The fabrication of patchy particles, summarized in Fig. 1a, starts with cross-linked amidinated polystyrene microspheres, 540 nm or 850 nm in diameter[30]. Small clusters of these microspheres are assembled using an emulsion-evaporation method[5] that produces so-called 'minimal-moment' clusters with reproducible symmetries and configurations: spheres, dumbbells, triangles, tetrahedra, triangular dipyramids, octahedra and pentagonal dipyramids, for clusters of $n = 1$–7 particles (Fig. 1b).

Patchy particles are formed from the clusters using a two-stage swelling process followed by polymerization[31]. First, a low-molecular-mass, water-insoluble organic compound (1-chlorodecane) is introduced into the colloidal clusters that are suspended in water with surfactant (sodium dodecyl sulphate, SDS). Adding a small amount of acetone to the suspension aids in the transport of the 1-chlorodecane into the colloidal clusters. We also introduce an oil-soluble initiator, benzoyl peroxide (BPO), and 1,2-dichloroethane, which dissolves BPO and is miscible with 1-chlorodecane. Subsequent stripping of the acetone and 1,2-dichloroethane from the solution traps both the 1-chlorodecane and the BPO in the polymer particles. The clusters are then swollen by the styrene monomer. The 1-chlorodecane introduced earlier acts as an osmotic swelling agent that increases the amount of monomer that can be absorbed by the clusters[32]. Because each cluster of a given number of particles contains the same amount of swelling agent, chemical equilibrium assures that

---

[1]Molecular Design Institute and Department of Chemistry, New York University, New York, New York 10003, USA. [2]The Dow Chemical Company, 2301 North Brazosport Boulevard, Freeport, Texas 77541, USA. [3]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA. [4]Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA. [5]Center for Soft Matter Research and Department of Physics, New York University, New York, New York 10003, USA.

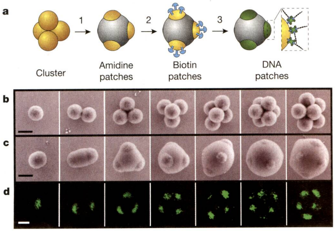


Figure 1 | DNA patchy particle fabrication. **a**, Preparation of colloidal particles with DNA-functionalized patches having well-defined symmetries. A four-patch particle is shown as an example. 1, A cluster of four amidinated polystyrene microspheres, prepared by the method of ref. 5, is swollen with styrene such that the extremities of the cluster—a tetrahedron in this case—protrude from the styrene droplet. The styrene is then polymerized and the protrusions from the original cluster become patches. 2, Biotin is site-specifically functionalized on the patches. 3, Biotinated DNA oligomers are introduced and bind to the particle patches via a biotin–streptavidin–biotin linkage. **b**, Electron micrographs of amidinated colloidal clusters, showing the particle configurations for clusters of $n = 1$–7 microspheres. **c**, Electron micrographs of amidinated patchy particles after encapsulation. The patches inherit the symmetries of their parent clusters. **d**, Confocal fluorescent images of corresponding patchy particles, verifying that only the patches are functionalized with DNA. The fluorescence comes from the dye-labelled streptavidin that links DNA with the patches. Scale bars, 500 nm.

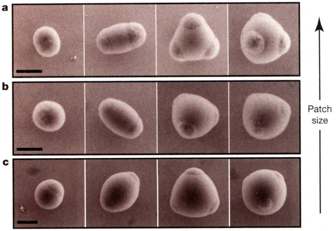


Figure 2 | Control of patch size. Electron micrographs of patchy particles, showing that the sizes of patches can be adjusted by changing encapsulation conditions. **a**, Particles with relatively large patches are fabricated when clusters are swollen with 1.0 ml of styrene monomer. Primary spheres are 540 nm in diameter. **b**, Under identical conditions, smaller patches are obtained when more monomer (1.2 ml) is added. **c**, Smaller patches, relative to particle size, are obtained using primary microspheres 850 nm in diameter. Using larger particles facilitates observation under an optical microscope. Divalent, trivalent and tetravalent particles from this batch were used in colloidal molecule formation, and the monovalent particles were used in kinetics study, as discussed below. Arrow indicates increasing patch size. Scale bars, 500 nm.

clusters of the same size all swell by the same amount, with the total amount of swelling controlled by the quantity of added monomer.

After swelling, we polymerize the styrene by thermally degrading the BPO previously introduced into each cluster. Swelling is controlled so that the extremities of the original clusters are not encapsulated, but are left as patches. Clusters of the same order *n* are encapsulated to the same extent, leading to uniform patch configurations, as seen in Fig. 1c, which shows scanning electron microscope (SEM) images of particles with 1 to 7 patches (see Supplementary Fig. 1a for higher-order patches). Using BPO as the initiator ensures that there are no functional groups introduced, so the surface created by swelling the clusters—the 'anti-patch' surface—is chemically inert and different from the patches: only the patches have the functional amidine groups.

Patch size is controlled during the swelling process by adjusting the amount of monomer that is introduced: the more monomer that is added, the smaller the patches are. Figure 2 shows that considerable variation in patch size can be achieved in this way. Small patches favour greater directionality, and larger patches permit multiple links per patch, as we discuss below.

A key design feature of our method is the use of clusters as intermediates. Their diversity in particle number and symmetry is translated directly to the number and symmetry of the particle patches. In contrast to the planar symmetry of Janus particles[23,33], the symmetries of these patchy particles are fully three-dimensional.

Our method converts essentially all the starting colloidal particles into particles with one or more patches. Each sample produced contains large scalable quantities of particles having different valence (numbers of patches), the relative distribution of which can be changed by adjusting the emulsification conditions used when making the clusters[34]. Using a higher shear rate, for example, makes smaller emulsion droplets, which skews the distribution towards lower-valence particles. We fractionate the particles through density gradient centrifugation, obtaining up to 12 clear bands corresponding to particles with different valence (see Supplementary Fig. 1b, c). Table 1 summarizes the fraction of particles obtained in each band for two different shearing conditions. For the lower-shear preparation, each of the four upper bands, which correspond to particles with 1 to 4 patches, contains 10^8 to 10^9 identical particles. For the higher-shear preparation, greater quantities are produced in the upper bands and lower quantities are produced in the lower bands. In most cases, we use conditions (see Methods) that produce the most 2-, 3- and 4-patch particles, which are the ones most useful for making analogues of common molecules. If pure samples containing patchy particles of identical valence are desired, then it is the fractionation step that ultimately limits the quantity available. Typically, we collect the same valence from up to 40 separations run in parallel, accumulating 10^9 to 10^{11} particles.

The amidine groups on the colloid surface are crucial to the patchy particle fabrication process. First, the positive charge created from the dissociation of amidine hydrogen chloride salt ($-C(NH)NH_3Cl$), along with the SDS surfactant, stabilizes the microspheres as well as the clusters by electrostatic repulsion. Second, when the clusters are swollen and encapsulated, the positive charges make the patches of the cluster more hydrophilic than the monomer–water interface, which is stabilized only by SDS. This difference in interfacial energies leads to finite contact angles and well-defined patches. Most importantly, the amidine groups can be easily functionalized in aqueous solution.

We functionalize the amidinated patches with biotin, and then use a biotin–streptavidin–biotin linkage to attach DNA with

Table 1 | Quantities of particles in the different bands

Number of patches	1	2	3	4	5	6	7
High shear	61%	15%	4%	1%	0.2%*	0.02%*	0.001%*
Low shear	7%	16%	25%	15%	8%	5%	3%

Density gradient centrifugation is used to fractionate the patchy particles. The fraction of identical particles obtained from a single centrifuge tube is shown. Fractions of 10%–20% correspond to 10^8–10^9 particles in a single fractionation.

* For these higher-valence particles, the fractions were estimated from their number ratio relative to lower-valence particles observed under a microscope (see Methods).

single-stranded 'sticky' ends. The first step uses sulpho-NHS-biotin (biotinamidohexanoic acid 3-sulpho-*N*-hydroxysuccinimide ester sodium salt, a water-soluble biotin derivative) and is carried out in phosphate-buffered saline (PBS) ($\mathrm{pH} = 7.42$), where the *N*-hydroxysuccinimide ester (NHS) can react with amidine groups and covalently link the biotin to the patches.

The DNA oligomer is prepared separately. It has three parts. At the 5′ end, it has a biotin as an anchoring molecule. In the middle, there is a 49-base-pair double helix that acts as a spacer. Finally, at the 3′ terminus, a single strand of 11 complementary or 8 palindrome base pairs forms the sticky end (for sequences, see Methods). Mixing DNA with streptavidin in a 1:1 ratio yields a streptavidin–DNA complex, which is then added to the biotin-functionalized patchy particles to produce DNA-functionalized patchy particles. The streptavidin contains a fluorescent tag for visualization by confocal microscopy. Figure 1d shows that only the patches of the particle are fluorescent, indicating that the streptavidin–DNA complex successfully coats the particle patches and that the amidine–NHS chemistry used for biotin functionalization works as designed.

The binding between patches on different particles is realized by hybridizing DNA oligomers on different patches. The oligomers, about 18 nm in length, provide short-range attractions and thus enforce the directionality defined by the particle patches. DNA is widely used for linking nanoparticles because it can be synthesized with control over the length and sequences of the base pairs, which, in turn, controls the specificity and the strength of interaction[35–38]. Hybridization of the complementary strands is fully reversible with temperature so that particle assembly can be controlled by varying temperature.

The oligomers we use to functionalize purified patchy particles are complementary DNA strands designated R (red) or G (green) and designed to bind selectively only to each other, or a palindrome P strand that only binds to other P strands. To differentiate the particles under the confocal microscope, red fluorescent (Alexa 647) streptavidin is used with R particles, and green fluorescent (Alexa 488) streptavidin is used with G particles (see Supplementary Fig. 2).

Assembly of colloidal molecules

With our collection of R, G and P patchy particles, we can build colloidal assemblies that mimic not only the geometry, but also the chemistry of molecules. Figure 3a (left panel) shows the formation of AB-type colloidal molecules from two 1-patch particles with complementary sticky ends. The system produces colloidal dumbbells without the random aggregation observed using spherical particles uniformly coated with DNA, and consistently with there being only one patch per particle. The confocal fluorescent image in Fig. 3a (middle panel) shows only complementary R–G particle pairs and no R–R or G–G pairs, confirming that DNA hybridization drives particle assembly. The resulting dumbbells are the colloidal analogues of AB-type molecules such as hydrogen chloride (Fig. 3a, right panel). Here, in contrast to hydrogen and chlorine, the sizes of the two atoms are the same, although they need not be. Patchy particles of different sizes can be fabricated and DNA bonds of various strengths can be used, so colloidal molecules of different size ratio and bond strength can be obtained.

Figure 3b shows linear AB_2-type colloidal molecules, the colloidal analogues of molecules like carbon dioxide (CO_2) or beryllium chloride ($BeCl_2$), that are obtained when green divalent (2-patch) particles are mixed with red monovalent particles. Triangle-like AB_3 (Fig. 3c) and tetrahedron-like AB_4 (Fig. 3d) colloidal molecules are similarly obtained by mixing trivalent (3-patch) particles and tetravalent (4-patch) particles, respectively, with monovalent particles (see Supplementary Video 1 for all colloidal molecules).

Bonding specificity is critical for the formation of all of the AB_n structures. It promotes the formation of AB bonds while prohibiting the formation of AA and BB bonds, ensuring that the divalent, trivalent and tetravalent particles can act as the central atoms and the bonding interactions mimic that of atomic orbitals in geometry and valence. Complementary monovalent particles then serve as ligands that form bonds with the central atom. The confocal images in Fig. 3 (middle panels) show the directionality and specificity of the interactions between the central atoms and their monovalent particle ligands.

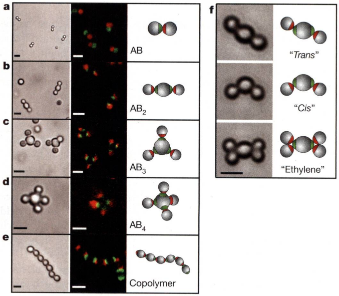


Figure 3 | Specific directional bonding between colloidal atoms observed with optical microscopes. **a**–**e**, Bright-field (left panels), confocal fluorescent (middle panels), and schematic images (right panels), show colloidal molecules self-assembled from patchy particles. **a**, Complementary green and red monovalent particles form dumbbell-shaped AB-type molecules. Supracolloidal molecules AB_2, AB_3 and AB_4 are formed by mixing red monovalent with green divalent (**b**), trivalent (**c**) and tetravalent (**d**) particles. **e**, If complementary divalent particles are mixed, linear alternating polymer chains spontaneously assemble. **f**, When particles with bigger patches are used, *cis*–*trans*-like isomers can form. Introducing more monovalent particles leads to ethylene-like colloidal molecules. Images are bright-field (left panels) and schematic (right panels). Scale bars, 2 μm.

Other structures that can be made include colloidal analogues of alternating copolymers, formed using complementary divalent particles. Figure 3e shows that only green and red divalent particles bind to each other.

If particles have patches big enough to accommodate more than one complementary particle, molecular isomers and branched polymers are obtained. Figure 3f shows two isomers of a nonlinear AB_2-type assembly that mimic the *cis*- and *trans*-conformations of molecules with a double bond. Such isomers may behave quite differently from one another in diffusion, rotation and reactivity. Additional monovalent particles can bind to the isomers and form ethylene-like structures (Fig. 3f, bottom panels). In the assembly of colloidal polymers from divalent particles, particles with bigger patches lead to branched chains and cross-linked networks (see Supplementary Fig. 3a). These results highlight the importance of controlling the patch size and in particular the ability to make patches sufficiently small that steric hindrance prevents more than one particle from attaching (see Supplementary Fig. 4 for SEM pictures of patchy particles used in colloidal molecules and geometry analysis for hindrance).

Self-complementary palindrome strands can also be used for self-assembly of mono- and divalent particles. Monovalent particles with palindrome sticky DNA yield A_2-type colloidal molecules, analogous to H_2 or Cl_2, whereas divalent particles yield homopolymers. One can also envision higher-order palindrome particles that might assemble into extended open structures like a diamond lattice[39].

Colloidal reactions

The self-assembly of patchy particles into a target structure can be viewed as a 'colloidal reaction' or more generally as "supracolloidal chemistry"[40]. As in conventional chemical reactions, colloidal particles with a particular morphology and binding capacity are used as reagents and mixed together stoichiometrically. For example, four equivalents of monovalent and one equivalent of complementary tetravalent particles assemble into AB_4 colloidal molecules. The yield is about 50% after a few days; that is, half of the tetravalent particles have four monovalent particles attached, with the remainder consisting of incomplete structures like AB_3, AB_2 and AB. Using an excess of the monovalent particles increases the yield of the final AB_4 product, just as for conventional chemical reactions. A fivefold excess of monovalent particles, for example, increases the yield of AB_4 to 80% (see Supplementary Fig. 3b). It should also be possible to increase the yield by increasing the strength of the DNA-mediated patch binding, which can be done by increasing the length of the DNA sticky ends or by increasing the density of the DNA attached to the patches[41].

An obvious and important difference between the molecular and colloidal domains is the size of the constituents. The much larger colloids exhibit much slower dynamics and reaction kinetics, and thus can be observed *in situ* under an optical microscope. As shown in Fig. 4a, the formation of an AB_4 molecule proceeds by the central tetravalent particle picking up monovalent particles, one at a time, over a period of about 30 min (see Supplementary Video 2). In the case of divalent particle chain formation, the 'polymerization' also follows a step-growth mechanism. Figure 4b illustrates how a polymer chain can be extended by adding divalent particles one by one at the end (see Supplementary Video 3). Alternatively, two polymer chains can fuse into a longer chain.

We can understand the stepwise growth mechanism by examining the kinetics of formation of the AB_3 molecules (see Supplementary Fig. 5 and Supplementary Video 4). We first heat a trivalent and monovalent particle mixture, with monovalent particles in large excess, to 55 °C, well above the melting temperature T_m (50 °C) of the DNA, thus causing the particles to dissociate. The system is then quenched to room temperature, well below T_m, so that the reaction kinetics is controlled by diffusion and by the size of the sticky patches. The collision frequency between monovalent and trivalent particles can be estimated from the well-known Smoluchowski equation[42]:

$$J \approx 4\pi(b_m + b_t)(D_m + D_t)C_m$$

where $b_m = 0.49$ µm is the radius of the monovalent particle, $b_t = 0.91$ µm is the radius of the trivalent particle, and C_m is the number concentration of the monovalent particles, estimated to be one particle per 50 μm^3 from direct observation. The diffusion coefficients are about $D_m = 0.50\ \mu m^2\ s^{-1}$ and $D_t = 0.28\ \mu m^2\ s^{-1}$ for the monovalent and trivalent particles, respectively. These values give a collision frequency between trivalent and monovalent particles of $0.27\ s^{-1}$ or a time between collisions of 3.6 s. Not every collision results in a bond, however, owing to the anisotropic nature of patchy particles. Only collisions between patches with complementary DNA can result in adhesion, with a rate proportional to the patch size. Thus, smaller patches form bonds more slowly than larger patches.

We define A as the fractional surface area of a particle occupied by patches (see Supplementary Figs 4 and 5), with values estimated from SEM measurements of $A_m = 0.23$ for the monovalent particle and $A_t = 0.077$ for all three patches of the trivalent particle (see Supplementary Table 1). The estimated reaction time for the first monovalent particle to adhere to a trivalent particle is $1/JA_mA_t$ or about 3.4 min. The area A for the trivalent–monovalent particle assembly immediately falls to 0.040 because one of the three patches is covered, and because the attached monovalent particle increases the total surface area of the complex. With two monovalent particles attached, A falls to 0.016. Thus the times for the second and third monovalent particles to attach are estimated to be 6.5 min and 16 min, respectively, which is consistent with the times observed experimentally and with the observed stepwise assembly of patchy particles (see also Supplementary Video 2).

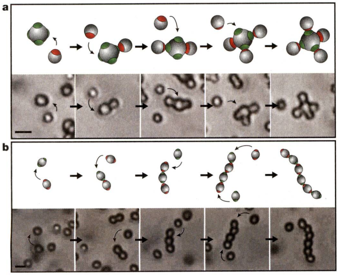


Figure 4 | Step-wise sequential kinetics of supracolloidal reactions. Schematic images (top panels) and snapshots from videos (bottom panels) show step-by-step reactions between colloidal atoms. Bent arrows point from the colloidal atom to the site where it is going to attach. Straight arrows indicate the time sequence. **a**, Monovalent particles attach to tetravalent particle, one by one, forming an AB_4-type colloidal molecule. **b**, Complementary divalent particles polymerize into a linear chain structure. Scale bar, 2 µm. Videos of these processes are available in the Supplementary Information.

Discussion

We expect that our colloids with valence can assemble not only into the molecular analogues shown here, but also into bulk colloidal phases of fundamental and practical interest. Tetrahedrally coordinated glasses, diamond crystals and empty liquids[43] have all proved difficult or impossible to make with existing colloids but should be accessible using our method. However, making scalable quantities of purified divalent and higher-valence colloidal atoms remains a challenge, owing to the limitations of fractionation by density gradient centrifugation. This difficulty might be overcome by large-scale separations[44], or by developing methods that produce clusters with controlled morphology so that no separation is needed. Indeed, the swelling and functionalization methods we use to make our colloidal particles could readily be adapted for use with clusters made using other recently developed techniques[22,25,45].

The ability to design colloidal particles with a variety of well-controlled three-dimensional bonding symmetries opens a wide spectrum of new structures for colloidal self-assembly, beyond colloidal assemblies whose structures are defined primarily by repulsive interactions and colloidal shape. Furthermore, the specificity of DNA interactions between patches means that colloids with different properties, such as size, colour, chemical functionality or electrical conductivity could be linked in well-defined sequences and orientations to make new functional materials. Such materials might include photonic crystals with programmed distributions of defects or three-dimensional electrically wired networks. We note that although valence and interaction strength in atomic systems are coupled by underlying quantum mechanical rules, they can be independently varied in our system. This raises the intriguing prospect of our patchy particles not merely mimicking atoms, but functioning as "designer atoms"[46] that can undergo reactions, unconstrained by the rules that govern bonding at the atomic scale, to yield structures with no analogues in the atomic or molecular realm[43,47].

METHODS SUMMARY

Cross-linked amidinated polystyrene microspheres were synthesized using a surfactant-free emulsion polymerization method. The amidinated clusters were prepared as described by ref. 5. Shearing conditions were optimized to control cluster distribution. A two-stage swelling and polymerization method was used to encapsulate the clusters, thereby fabricating mixtures of patchy particles that were separated by density gradient centrifugation. After purification, the separated particles were dispersed in 10 mM PBS (pH 7.42, 100 mM NaCl) containing 0.1% (w/w) Triton X-100 and reacted with sulpho-NHS-biotin to convert the functionalities on the patches from amidines to biotins. 5′-Biotin-DNA was mixed with fluorescent streptavidin in a 1:1 ratio and the resulting complex was used to attach DNA to the biotinylated patches. Finally, the DNA-containing particles were washed with and dispersed in an aqueous PBS containing 1% (w/w) Pluronic F127. This suspension was used for all self-assembly experiments. For the self-assembly studies, the mixture of interest was sealed in a hydrophobic capillary tube and imaged using optical microscopy. Particles dried on a silicon wafer were imaged by field-emission SEM. The fluorescent images were obtained using a Leica SP5 confocal fluorescence microscope. Laser lines at 488 nm and 633 nm were used to excite the green and red fluorescence.

Full Methods and any associated references are available in the online version of the paper.

Received 12 June; accepted 4 September 2012.

1. Glotzer, S. C. & Solomon, M. J. Anisotropy of building blocks and their assembly into complex structures. *Nature Mater.* **6,** 557–562 (2007).
2. Sacanna, S. & Pine, D. J. Shape-anisotropic colloids: building blocks for complex assemblies. *Curr. Opin. Colloid Interf. Sci.* **16,** 96–105 (2011).
3. Rossi, L. *et al.* Cubic crystals from cubic colloids. *Soft Matter* **7,** 4139–4142 (2011).
4. Meng, G., Arkus, N., Brenner, M. P. & Manoharan, V. N. The free-energy landscape of clusters of attractive hard spheres. *Science* **327,** 560–563 (2010).
5. Manoharan, V. N., Elsesser, M. T. & Pine, D. J. Dense packing and symmetry in small clusters of microspheres. *Science* **301,** 483–487 (2003).
6. Kraft, D. J. *et al.* Self-assembly of colloids with liquid protrusions. *J. Am. Chem. Soc.* **131,** 1182–1186 (2009).
7. Sacanna, S., Irvine, W. T. M., Chaikin, P. M. & Pine, D. J. Lock and key colloids. *Nature* **464,** 575–578 (2010).
8. Zoldesi, C. I., van Walree, C. A. & Imhof, A. Deformable hollow hybrid silica/siloxane colloids by emulsion templating. *Langmuir* **22,** 4343–4352 (2006).
9. Li, F., Josephson, D. P. & Stein, A. Colloidal assembly: the road from particles to colloidal molecules and crystals. *Angew. Chem. Int. Edn* **50,** 360–388 (2011).
10. Leunissen, M. E. *et al.* Ionic colloidal crystals of oppositely charged particles. *Nature* **437,** 235–240 (2005).
11. Macfarlane, R. J. *et al.* Nanoparticle superlattice engineering with DNA. *Science* **334,** 204–208 (2011).
12. Velikov, K. P., Christova, C. G., Dullens, R. P. A. & van Blaaderen, A. Layer-by-layer growth of binary colloidal crystals. *Science* **296,** 106–109 (2002).
13. Ho, K. M., Chan, C. T. & Soukoulis, C. M. Existence of a photonic gap in periodic dielectric structures. *Phys. Rev. Lett.* **65,** 3152–3155 (1990).
14. Nelson, D. R. Toward a tetravalent chemistry of colloids. *Nano Lett.* **2,** 1125–1129 (2002).
15. Glotzer, S. C. Some assembly required. *Science* **306,** 419–420 (2004).
16. Manoharan, V. N. *Building Materials with Colloidal Spheres.* PhD thesis, Univ. California, Santa Barbara (2004).
17. Breed, D. R. *Engineered Colloids: Patchy Particles with Reversible, Directional Interactions.* PhD thesis, Univ. California, Santa Barbara (2007).
18. Bianchi, E., Blaak, R. & Likos, C. N. Patchy colloids: state of the art and perspectives. *Phys. Chem. Chem. Phys.* **13,** 6397–6410 (2011).
19. Zhang, Z. & Glotzer, S. C. Self-assembly of patchy particles. *Nano Lett.* **4,** 1407–1413 (2004).
20. Xu, X., Rosi, N. L., Wang, Y., Huo, F. & Mirkin, C. A. Asymmetric functionalization of gold nanoparticles with oligonucleotides. *J. Am. Chem. Soc.* **128,** 9286–9287 (2006).
21. Huo, F., Lytton-Jean, A. K. R. & Mirkin, C. A. Asymmetric functionalization of nanoparticles based on thermally addressable DNA interconnects. *Adv. Mater.* **18,** 2304–2306 (2006).
22. Hong, L., Cacciuto, A., Luijten, E. & Granick, S. Clusters of amphiphilic colloidal spheres. *Langmuir* **24,** 621–625 (2008).
23. Chen, Q., Bae, S. C. & Granick, S. Directed self-assembly of a colloidal kagome lattice. *Nature* **469,** 381–384 (2011).
24. Jones, M. R., Macfarlane, R. J., Prigodich, A. E., Patel, P. C. & Mirkin, C. A. Nanoparticle shape anisotropy dictates the collective behavior of surface-bound ligands. *J. Am. Chem. Soc.* **133,** 18865–18869 (2011).
25. Kraft, D. J., Groenewold, J. & Kegel, W. K. Colloidal molecules with well-controlled bond angles. *Soft Matter* **5,** 3823–3826 (2009).
26. Cho, Y.-S. *et al.* Particles with coordinated patches or windows from oil-in-water emulsions. *Chem. Mater.* **19,** 3183–3193 (2007).
27. Zhang, G., Wang, D. & Möhwald, H. Patterning microsphere surfaces by templating colloidal crystals. *Nano Lett.* **5,** 143–146 (2005).
28. Pawar, A. B. & Kretzschmar, I. Multifunctional patchy particles by glancing angle deposition. *Langmuir* **25,** 9057–9063 (2009).
29. Pawar, A. B. & Kretzschmar, I. Fabrication, assembly, and application of patchy particles. *Macromol. Rapid Commun.* **31,** 150–168 (2010).
30. Ottewill, R. & Shaw, J. Studies on the preparation and characterization of monodisperse polystyrene lattices. *Colloid Polym. Sci.* **215,** 161–166 (1967).
31. Ugelstad, J., Kaggerud, K. H., Hansen, F. K. & Berge, A. Absorption of low molecular weight compounds in aqueous dispersions of polymer-oligomer particles, 2: A two step swelling process of polymer particles giving an enormous increase in absorption capacity. *Makromol. Chem.* **180,** 737–744 (1979).
32. Ugelstad, J. *et al.* Thermodynamics of swelling. Preparation and application of some composite, monosized polymer particles. *Makromol. Chem.* **10,** 215–234 (1985).
33. Walther, A. & Muller, A. H. E. Janus particles. *Soft Matter* **4,** 663–668 (2008).
34. Zerrouki, D. *et al.* Preparation of doublet, triangular, and tetrahedral colloidal clusters by controlled emulsification. *Langmuir* **22,** 57–62 (2006).
35. Mirkin, C. A., Letsinger, R. L., Mucic, R. C. & Storhoff, J. J. A DNA-based method for rationally assembling nanoparticles into macroscopic materials. *Nature* **382,** 607–609 (1996).
36. Alivisatos, A. P. *et al.* Organization of 'nanocrystal molecules' using DNA. *Nature* **382,** 609–611 (1996).
37. Leunissen, M. E. *et al.* Switchable self-protected attractions in DNA-functionalized colloids. *Nature Mater.* **8,** 590–595 (2009).
38. Nykypanchuk, D., Maye, M. M., van der Lelie, D. & Gang, O. DNA-guided crystallization of colloidal nanoparticles. *Nature* **451,** 549–552 (2008).
39. Zhang, Z., Keys, A. S., Chen, T. & Glotzer, S. C. Self-assembly of patchy particles into diamond structures through molecular mimicry. *Langmuir* **21,** 11547–11551 (2005).
40. Chen, Q. *et al.* Supracolloidal reaction kinetics of Janus spheres. *Science* **331,** 199–202 (2011).
41. Dreyfus, R. *et al.* Aggregation-disaggregation transition of DNA-coated colloids: Experiments and theory. *Phys. Rev. E* **81,** 041404 (2010).
42. van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* 3rd edn, 292–297 (Elsevier, 2007).
43. Romano, F., Sanz, E. & Sciortino, F. Crystallization of tetrahedral patchy particles in silico. *J. Chem. Phys.* **134,** 174502–174508 (2011).
44. Johnson, P. M., van Kats, C. M. & van Blaaderen, A. Synthesis of colloidal silica dumbbells. *Langmuir* **21,** 11510–11517 (2005).
45. Perro, A. *et al.* A chemical synthetic route towards "colloidal molecules". *Angew. Chem. Int. Edn* **48,** 361–365 (2009).
46. Frenkel, D. Playing tricks with designer "atoms". *Science* **296,** 65–66 (2002).
47. Miller, M. A. & Wales, D. J. Novel structural motifs in clusters of dipolar spheres: knots, links, and coils. *J. Phys. Chem. B* **109,** 23109–23112 (2005).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. T. Elsesser, D. Kraft, and G.-R. Yi for discussions. This work was supported partially by the MRSEC Program of the National Science Foundation under award number DMR-0820341. Additional financial support was provided by the National Science Foundation (ChE-0911460). We acknowledge support from the MRI programme of the National Science Foundation under award number DMR-0923251 for the purchase of a Zeiss field emission scanning electron microscope.

Author Contributions V.N.M. and D.J.P. thought of patchy particles and how to make them. D.R.B. developed the functionalization of patches. Yufeng Wang and Yu Wang optimized the control of patch size and DNA functionalization, performed the experiments, and collected and analysed the data. A.D.H. helped with particle synthesis. L.F. designed the DNA sequences and modelled the kinetics. M.W. and D.J.P. supervised the project. D.J.P., Yufeng Wang and Yu Wang wrote the paper with revisions from all of the authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.J.P. (pine@nyu.edu) or M.W. (marcus.weck@nyu.edu).

doi:10.1038/nature11632

An integrated map of genetic variation from 1,092 human genomes

The 1000 Genomes Project Consortium*

By characterizing the geographic and functional spectrum of human genetic variation, the 1000 Genomes Project aims to build a resource to help to understand the genetic contribution to disease. Here we describe the genomes of 1,092 individuals from 14 populations, constructed using a combination of low-coverage whole-genome and exome sequencing. By developing methods to integrate information across several algorithms and diverse data sources, we provide a validated haplotype map of 38 million single nucleotide polymorphisms, 1.4 million short insertions and deletions, and more than 14,000 larger deletions. We show that individuals from different populations carry different profiles of rare and common variants, and that low-frequency variants show substantial geographic differentiation, which is further increased by the action of purifying selection. We show that evolutionary conservation and coding consequence are key determinants of the strength of purifying selection, that rare-variant load varies substantially across biological pathways, and that each individual contains hundreds of rare non-coding variants at conserved sites, such as motif-disrupting changes in transcription-factor-binding sites. This resource, which captures up to 98% of accessible single nucleotide polymorphisms at a frequency of 1% in related populations, enables analysis of common and low-frequency variants in individuals from diverse, including admixed, populations.

Recent efforts to map human genetic variation by sequencing exomes[1] and whole genomes[2–4] have characterized the vast majority of common single nucleotide polymorphisms (SNPs) and many structural variants across the genome. However, although more than 95% of common (>5% frequency) variants were discovered in the pilot phase of the 1000 Genomes Project, lower-frequency variants, particularly those outside the coding exome, remain poorly characterized. Low-frequency variants are enriched for potentially functional mutations, for example, protein-changing variants, under weak purifying selection[1,5,6]. Furthermore, because low-frequency variants tend to be recent in origin, they exhibit increased levels of population differentiation[6–8]. Characterizing such variants, for both point mutations and structural changes, across a range of populations is thus likely to identify many variants of functional importance and is crucial for interpreting individual genome sequences, to help separate shared variants from those private to families, for example.

We now report on the genomes of 1,092 individuals sampled from 14 populations drawn from Europe, East Asia, sub-Saharan Africa and the Americas (Supplementary Figs 1 and 2), analysed through a combination of low-coverage (2–6×) whole-genome sequence data, targeted deep (50–100×) exome sequence data and dense SNP genotype data (Table 1 and Supplementary Tables 1–3). This design was shown by the pilot phase[2] to be powerful and cost-effective in discovering and genotyping all but the rarest SNP and short insertion and deletion (indel) variants. Here, the approach was augmented with statistical methods for selecting higher quality variant calls from candidates obtained using multiple algorithms, and to integrate SNP, indel and larger structural variants within a single framework (see

Table 1 | Summary of 1000 Genomes Project phase I data

	Autosomes	Chromosome X	GENCODE regions*
Samples	1,092	1,092	1,092
Total raw bases (Gb)	19,049	804	327
Mean mapped depth (×)	5.1	3.9	80.3
SNPs			
No. sites overall	36.7 M	1.3 M	498 K
Novelty rate†	58%	77%	50%
No. synonymous/non-synonymous/nonsense	NA	4.7/6.5/0.097 K	199/293/6.3 K
Average no. SNPs per sample	3.60 M	105 K	24.0 K
Indels			
No. sites overall	1.38 M	59 K	1,867
Novelty rate†	62%	73%	54%
No. inframe/frameshift	NA	19/14	719/1,066
Average no. indels per sample	344 K	13 K	440
Genotyped large deletions			
No. sites overall	13.8 K	432	847
Novelty rate†	54%	54%	50%
Average no. variants per sample	717	26	39

NA, not applicable.
* Autosomal genes only.
† Compared with dbSNP release 135 (Oct 2011), excluding contribution from phase I 1000 Genomes Project (or equivalent data for large deletions).

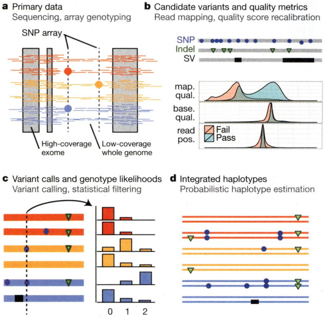
---

*Lists of participants and their affiliations appear at the end of the paper.

BOX 1

Constructing an integrated map of variation

The 1,092 haplotype-resolved genomes released as phase I by the 1000 Genomes Project are the result of integrating diverse data from multiple technologies generated by several centres between 2008 and 2010. The Box 1 Figure describes the process leading from primary data production to integrated haplotypes.


a Primary data
Sequencing, array genotyping
SNP array
High-coverage exome
Low-coverage whole genome
b Candidate variants and quality metrics
Read mapping, quality score recalibration
SNP
Indel
SV
map. qual.
base. qual.
read pos.
Fail
Pass
c Variant calls and genotype likelihoods
Variant calling, statistical filtering
0
1
2
d Integrated haplotypes
Probabilistic haplotype estimation

a, Unrelated individuals (see Supplementary Table 10 for exceptions) were sampled in groups of up to 100 from related populations (Wright's F_{ST} typically <1%) within broader geographical or ancestry-based groups[2]. Primary data generated for each sample consist of low-coverage (average 5×) whole-genome and high-coverage (average 80× across a consensus target of 24 Mb spanning more than 15,000 genes) exome sequence data, and high density SNP array information. **b,** Following read-alignment, multiple algorithms were used to identify candidate variants. For each variant, quality metrics were obtained, including information about the uniqueness of the surrounding sequence (for example, mapping quality (map. qual.)), the quality of evidence supporting the variant (for example, base quality (base. qual.) and the position of variant bases within reads (read pos.)), and the distribution of variant calls in the population (for example, inbreeding coefficient). Machine-learning approaches using this multidimensional information were trained on sets of high-quality known variants (for example, the high-density SNP array data), allowing variant sites to be ranked in confidence and subsequently thresholded to ensure low FDR. **c,** Genotype likelihoods were used to summarize the evidence for each genotype at bi-allelic sites (0, 1 or 2 copies of the variant) in each sample at every site. **d,** As the evidence for a single genotype is typically weak in the low-coverage data, and can be highly variable in the exome data, statistical methods were used to leverage information from patterns of linkage disequilibrium, allowing haplotypes (and genotypes) to be inferred.

Box 1 and Supplementary Fig. 1). Because of the challenges of identifying large and complex structural variants and shorter indels in regions of low complexity, we focused on conservative but high-quality subsets: biallelic indels and large deletions.

Overall, we discovered and genotyped 38 million SNPs, 1.4 million bi-allelic indels and 14,000 large deletions (Table 1). Several technologies were used to validate a frequency-matched set of sites to assess and control the false discovery rate (FDR) for all variant types. Where results were clear, 3 out of 185 exome sites (1.6%), 5 out of 281 low-coverage sites (1.8%) and 72 out of 3,415 large deletions (2.1%) could not be validated (Supplementary Information and Supplementary Tables 4–9). The initial indel call set was found to have a high FDR (27 out of 76), which led to the application of further filters, leaving an implied FDR of 5.4% (Supplementary Table 6 and Supplementary Information). Moreover, for 2.1% of low-coverage SNP and 18% of indel sites, we found inconsistent or ambiguous results, indicating that substantial challenges remain in characterizing variation in low-complexity genomic regions. We previously described the 'accessible genome': the fraction of the reference genome in which short-read data can lead to reliable variant discovery. Through longer read lengths, the fraction accessible has increased from 85% in the pilot phase to 94% (available as a genome annotation; see Supplementary Information), and 1.7 million low-quality SNPs from the pilot phase have been eliminated.

By comparison to external SNP and high-depth sequencing data, we estimate the power to detect SNPs present at a frequency of 1% in the study samples is 99.3% across the genome and 99.8% in the consensus exome target (Fig. 1a). Moreover, the power to detect SNPs at 0.1% frequency in the study is more than 90% in the exome and nearly 70% across the genome. The accuracy of individual genotype calls at heterozygous sites is more than 99% for common SNPs and 95% for SNPs at a frequency of 0.5% (Fig. 1b). By integrating linkage disequilibrium information, genotypes from low-coverage data are as accurate as those from high-depth exome data for SNPs with frequencies >1%. For very rare SNPs (≤0.1%, therefore present in one or two copies), there is no gain in genotype accuracy from incorporating linkage disequilibrium information and accuracy is lower. Variation among samples in genotype accuracy is primarily driven by sequencing depth (Supplementary Fig. 3) and technical issues such as sequencing platform and version (detectable by principal component analysis; Supplementary Fig. 4), rather than by population-level characteristics. The accuracy of inferred haplotypes at common SNPs was estimated by comparison to SNP data collected on mother–father–offspring trios for a subset of the samples. This indicates that a phasing (switch) error is made, on average, every 300–400 kilobases (kb) (Supplementary Fig. 5).

A key goal of the 1000 Genomes Project was to identify more than 95% of SNPs at 1% frequency in a broad set of populations. Our current resource includes ~50%, 98% and 99.7% of the SNPs with frequencies of ~0.1%, 1.0% and 5.0%, respectively, in ~2,500 UK-sampled genomes (the Wellcome Trust-funded UK10K project), thus

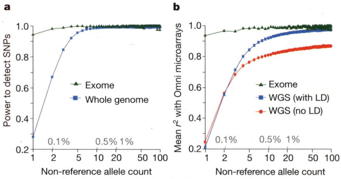


Figure 1 | Power and accuracy. a, Power to detect SNPs as a function of variant count (and proportion) across the entire set of samples, estimated by comparison to independent SNP array data in the exome (green) and whole genome (blue). **b,** Genotype accuracy compared with the same SNP array data as a function of variant frequency, summarized by the r^2 between true and inferred genotype (coded as 0, 1 and 2) within the exome (green), whole genome after haplotype integration (blue), and whole genome without haplotype integration (red). LD, linkage disequilibrium; WGS, whole-genome sequencing.

meeting this goal. However, coverage may be lower for populations not closely related to those studied. For example, our resource includes only 23.7%, 76.9% and 99.3% of the SNPs with frequencies of $\sim$0.1%, 1.0% and 5.0%, respectively, in $\sim$2,000 genomes sequenced in a study of the isolated population of Sardinia (the SardiNIA study).

Genetic variation within and between populations

The integrated data set provides a detailed view of variation across several populations (illustrated in Fig. 2a). Most common variants (94% of variants with frequency $\geq$5% in Fig. 2a) were known before the current phase of the project and had their haplotype structure mapped through earlier projects[2,9]. By contrast, only 62% of variants in the range 0.5–5% and 13% of variants with frequencies of $\leq$0.5% had been described previously. For analysis, populations are grouped by the predominant component of ancestry: Europe (CEU (see Fig. 2a for definitions of this and other populations), TSI, GBR, FIN and IBS), Africa (YRI, LWK and ASW), East Asia (CHB, JPT and CHS) and the Americas (MXL, CLM and PUR). Variants present at 10% and above across the entire sample are almost all found in all of the populations studied. By contrast, 17% of low-frequency variants in the range 0.5–5% were observed in a single ancestry group, and 53% of rare variants at 0.5% were observed in a single population (Fig. 2b). Within ancestry groups, common variants are weakly differentiated (most within-group estimates of Wright's fixation index (F_{ST}) are <1%; Supplementary Table 11), although below 0.5% frequency variants are up to twice as likely to be found within the same population compared with random samples from the ancestry group (Supplementary Fig. 6a). The degree of rare-variant differentiation varies between populations. For example, within Europe, the IBS and FIN populations carry excesses of rare variants (Supplementary Fig. 6b), which can arise through events such as recent bottlenecks[10], 'clan' breeding structures[11] and admixture with diverged populations[12].

Some common variants show strong differentiation between populations within ancestry-based groups (Supplementary Table 12), many of which are likely to have been driven by local adaptation either directly or through hitchhiking. For example, the strongest differentiation between African populations is within an NRSF (neuron-restrictive silencer factor) transcription-factor peak (PANC1 cell line)[13], upstream of *ST8SIA1* (difference in derived allele frequency LWK − YRI of 0.475 at rs7960970), whose product is involved in ganglioside generation[14]. Overall, we find a range of 17–343 SNPs (fewest = CEU − GBR, most = FIN − TSI) showing a difference in frequency of at least 0.25 between pairs of populations within an ancestry group.

The derived allele frequency distribution shows substantial divergence between populations below a frequency of 40% (Fig. 2c), such that individuals from populations with substantial African ancestry (YRI, LWK and ASW) carry up to three times as many low-frequency variants (0.5–5% frequency) as those of European or East Asian origin, reflecting ancestral bottlenecks in non-African populations[15]. However, individuals from all populations show an enrichment of rare variants (<0.5% frequency), reflecting recent explosive increases in population size and the effects of geographic differentiation[6,16]. Compared with the expectations from a model of constant population size, individuals from all populations show a substantial excess of high-frequency-derived variants (>80% frequency).

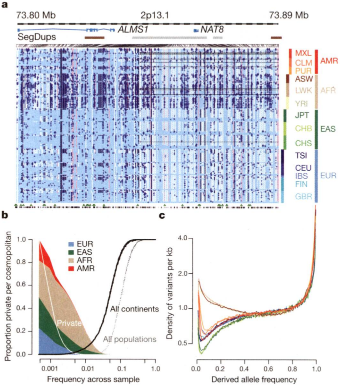


Figure 2 | The distribution of rare and common variants. **a**, Summary of inferred haplotypes across a 100-kb region of chromosome 2 spanning the genes *ALMS1* and *NAT8*, variation in which has been associated with kidney disease[45]. Each row represents an estimated haplotype, with the population of origin indicated on the right. Reference alleles are indicated by the light blue background. Variants (non-reference alleles) above 0.5% frequency are indicated by pink (typed on the high-density SNP array), white (previously known) and dark blue (not previously known). Low frequency variants (<0.5%) are indicated by blue crosses. Indels are indicated by green triangles and novel variants by dashes below. A large, low-frequency deletion (black line) spanning *NAT8* is present in some populations. Multiple structural haplotypes mediated by segmental duplications are present at this locus, including copy number gains, which were not genotyped for this study. Within each population, haplotypes are ordered by total variant count across the region. Population abbreviations: ASW, people with African ancestry in Southwest United States; CEU, Utah residents with ancestry from Northern and Western Europe; CHB, Han Chinese in Beijing, China; CHS, Han Chinese South, China; CLM, Colombians in Medellin, Colombia; FIN, Finnish in Finland; GBR, British from England and Scotland, UK; IBS, Iberian populations in Spain; LWK, Luhya in Webuye, Kenya; JPT, Japanese in Tokyo, Japan; MXL, people with Mexican ancestry in Los Angeles, California; PUR, Puerto Ricans in Puerto Rico; TSI, Toscani in Italia; YRI, Yoruba in Ibadan, Nigeria. Ancestry-based groups: AFR, African; AMR, Americas; EAS, East Asian; EUR, European. **b**, The fraction of variants identified across the project that are found in only one population (white line), are restricted to a single ancestry-based group (defined as in **a**, solid colour), are found in all groups (solid black line) and all populations (dotted black line). **c**, The density of the expected number of variants per kilobase carried by a genome drawn from each population, as a function of variant frequency (see Supplementary Information). Colours as in **a**. Under a model of constant population size, the expected density is constant across the frequency spectrum.

Because rare variants are typically recent, their patterns of sharing can reveal aspects of population history. Variants present twice across the entire sample (referred to as f_2 variants), typically the most recent of informative mutations, are found within the same population in 53% of cases (Fig. 3a). However, between-population sharing identifies recent historical connections. For example, if one of the individuals carrying an f_2 variant is from the Spanish population (IBS) and the other is not (referred to as IBS−X), the other individual is more likely to come from the Americas populations (48%, correcting for sample size) than from elsewhere in Europe (41%). Within the East Asian populations, CHS and CHB show stronger f_2 sharing to each other (58% and 53% of CHS−X and CHB−X variants, respectively) than either does to JPT, but JPT is closer to CHB than to CHS (44% versus 35% of JPT−X variants). Within African-ancestry populations, the ASW are closer to the YRI (42% of ASW−X f_2 variants) than the LWK (28%), in line with historical information[17] and genetic evidence based on common SNPs[18]. Some sharing patterns are surprising; for example, 2.5% of the f_2 FIN−X variants are shared with YRI or LWK populations.

Independent evidence about variant age comes from the length of the shared haplotypes on which they are found. We find, as expected,

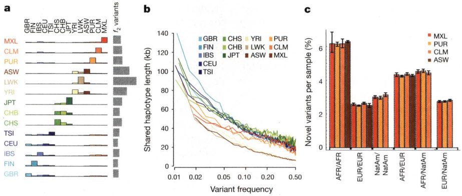


Figure 3 | Allele sharing within and between populations. a, Sharing of f_2 variants, those found exactly twice across the entire sample, within and between populations. Each row represents the distribution across populations for the origin of samples sharing an f_2 variant with the target population (indicated by the left-hand side). The grey bars represent the average number of f_2 variants carried by a randomly chosen genome in each population. **b**, Median length of haplotype identity (excluding cryptically related samples and singleton variants, and allowing for up to two genotype errors) between two chromosomes that share variants of a given frequency in each population. Estimates are from 200 randomly sampled regions of 1 Mb each and up to 15 pairs of individuals for each variant. **c**, The average proportion of variants that are new (compared with the pilot phase of the project) among those found in regions inferred to have different ancestries within ASW, PUR, CLM and MXL populations. Error bars represent 95% bootstrap confidence intervals. NatAm, Native American.

a negative correlation between variant frequency and the median length of shared haplotypes, such that chromosomes carrying variants at 1% frequency share haplotypes of 100–150 kb (typically 0.08–0.13 cM; Fig. 3b and Supplementary Fig. 7a), although the distribution is highly skewed and 2–5% of haplotypes around the rarest SNPs extend over 1 megabase (Mb) (Supplementary Fig. 7b, c). Haplotype phasing and genotype calling errors will limit the ability to detect long shared haplotypes, and the observed lengths are a factor of 2–3 times shorter than predicted by models that allow for recent explosive growth[6] (Supplementary Fig. 7a). Nevertheless, the haplotype length for variants shared within and between populations is informative about relative allele age. Within populations and between populations in which there is recent shared ancestry (for example, through admixture and within continents), f_2 variants typically lie on long shared haplotypes (median within ancestry group 103 kb; Supplementary Fig. 8). By contrast, between populations with no recent shared ancestry, f_2 variants are present on very short haplotypes, for example, an average of 11 kb for FIN − YRI f_2 variants (median between ancestry groups excluding admixture is 15 kb), and are therefore likely to reflect recurrent mutations and chance ancient coalescent events.

To analyse populations with substantial historical admixture, statistical methods were applied to each individual to infer regions of the genome with different ancestries. Populations and individuals vary substantially in admixture proportions. For example, the MXL population contains the greatest proportion of Native American ancestry (47% on average compared with 24% in CLM and 13% in PUR), but the proportion varies from 3% to 92% between individuals (Supplementary Fig. 9a). Rates of variant discovery, the ratio of non-synonymous to synonymous variation and the proportion of variants that are new vary systematically between regions with different ancestries. Regions of Native American ancestry show less variation, but a higher fraction of the variants discovered are novel (3.0% of variants per sample; Fig. 3c) compared with regions of European ancestry (2.6%). Regions of African ancestry show the highest rates of novelty (6.2%) and heterozygosity (Supplementary Fig. 9b, c).

The functional spectrum of human variation

The phase I data enable us to compare, for different genomic features and variant types, the effects of purifying selection on evolutionary conservation[19], the allele frequency distribution and the level of differentiation between populations. At the most highly conserved coding sites, 85% of non-synonymous variants and more than 90% of stop-gain and splice-disrupting variants are below 0.5% in frequency, compared with 65% of synonymous variants (Fig. 4a). In general, the rare variant excess tracks the level of evolutionary conservation for variants of most functional consequence, but varies systematically between types (for example, for a given level of conservation enhancer variants have a higher rare variant excess than variants in transcription-factor motifs). However, stop-gain variants and, to a lesser extent, splice-site disrupting changes, show increased rare-variant excess whatever the conservation of the base in which they occur, as such mutations can be highly deleterious whatever the level of sequence conservation. Interestingly, the least conserved splice-disrupting variants show similar rare-variant loads to synonymous and non-coding regions, suggesting that these alternative transcripts are under very weak selective constraint. Sites at which variants are observed are typically less conserved than average (for example, sites with non-synonymous variants are, on average, as conserved as third codon positions; Supplementary Fig. 10).

A simple way of estimating the segregating load arising from rare, deleterious mutations across a set of genes comes from comparing the

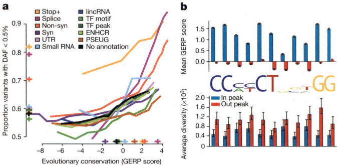


Figure 4 | Purifying selection within and between populations. a, The relationship between evolutionary conservation (measured by GERP score[19]) and rare variant proportion (fraction of all variants with derived allele frequency (DAF) $< 0.5\%$) for variants occurring in different functional elements and with different coding consequences. Crosses indicate the average GERP score at variant sites (x axis) and the proportion of rare variants (y axis) in each category. ENHCR, enhancer; lincRNA, large intergenic non-coding RNA; non-syn, non-synonymous; PSEUG, pseudogene; syn, synonymous; TF, transcription factor. **b**, Levels of evolutionary conservation (mean GERP score, top) and genetic diversity (per-nucleotide pairwise differences, bottom) for sequences matching the CTCF-binding motif within CTCF-binding peaks, as identified experimentally by ChIP-seq in the ENCODE project[13] (blue) and in a matched set of motifs outside peaks (red). The logo plot shows the distribution of identified motifs within peaks. Error bars represent ±2 s.e.m.

ratios of non-synonymous to synonymous variants in different frequency ranges. The non-synonymous to synonymous ratio among rare (<0.5%) variants is typically in the range 1–2, and among common variants in the range 0.5–1.5, suggesting that 25–50% of rare non-synonymous variants are deleterious. However, the segregating rare load among gene groups in KEGG pathways[20] varies substantially (Supplementary Fig. 11a and Supplementary Table 13). Certain groups (for example, those involving extracellular matrix (ECM)–receptor interactions, DNA replication and the pentose phosphate pathway) show a substantial excess of rare coding mutations, which is only weakly correlated with the average degree of evolutionary conservation. Pathways and processes showing an excess of rare functional variants vary between continents (Supplementary Fig. 11b). Moreover, the excess of rare non-synonymous variants is typically higher in populations of European and East Asian ancestry (for example, the ECM–receptor interaction pathway load is strongest in European populations). Other groups of genes (such as those associated with allograft rejection) have a high non-synonymous to synonymous ratio in common variants, potentially indicating the effects of positive selection.

Genome-wide data provide important insights into the rates of functional polymorphism in the non-coding genome. For example, we consider motifs matching the consensus for the transcriptional repressor CTCF, which has a well-characterized and highly conserved binding motif[21]. Within CTCF-binding peaks experimentally defined by chromatin-immunoprecipitation sequencing (ChIP-seq), the average levels of conservation within the motif are comparable to third codon positions, whereas there is no conservation outside peaks (Fig. 4b). Within peaks, levels of genetic diversity are typically reduced 25–75%, depending on the position in the motif (Fig. 4b). Unexpectedly, the reduction in diversity at some degenerate positions, for example, at position 8 in the motif, is as great as that at non-degenerate positions, suggesting that motif degeneracy may not have a simple relationship with functional importance. Variants within peaks show a weak but consistent excess of rare variation (proportion with frequency <0.5% is 61% within peaks compared with 58% outside peaks; Supplementary Fig. 12), supporting the hypothesis that regulatory sequences contain substantial amounts of weakly deleterious variation.

Purifying selection can also affect population differentiation if its strength and efficacy vary among populations. Although the magnitude of the effect is weak, non-synonymous variants consistently show greater levels of population differentiation than synonymous variants, for variants of frequencies of less than 10% (Supplementary Fig. 13).

Uses of 1000 Genomes Project data in medical genetics

Data from the 1000 Genomes Project are widely used to screen variants discovered in exome data from individuals with genetic disorders[22] and in cancer genome projects[23]. The enhanced catalogue presented here improves the power of such screening. Moreover, it provides a 'null expectation' for the number of rare, low-frequency and common variants with different functional consequences typically found in randomly sampled individuals from different populations.

Estimates of the overall numbers of variants with different sequence consequences are comparable to previous values[1,20–22] (Supplementary Table 14). However, only a fraction of these are likely to be functionally relevant. A more accurate picture of the number of functional variants is given by the number of variants segregating at conserved positions (here defined as sites with a genomic evolutionary rate profiling (GERP)[19] conservation score of >2), or where the function (for example, stop-gain variants) is strong and independent of conservation (Table 2). We find that individuals typically carry more than 2,500 non-synonymous variants at conserved positions, 20–40 variants identified as damaging[24] at conserved sites and about 150 loss-of-function (LOF) variants (stop-gains, frameshift indels in coding sequence and disruptions to essential splice sites). However, most of these are common (>5%) or low-frequency (0.5–5%), such that the numbers of rare (<0.5%) variants in these categories (which might be considered as pathological candidates) are much lower; 130–400 non-synonymous variants per individual, 10–20 LOF variants, 2–5 damaging mutations, and 1–2 variants identified previously from cancer genome sequencing[25]. By comparison with synonymous variants, we can estimate the excess of rare variants; those mutations that are sufficiently deleterious that they will never reach high frequency. We estimate that individuals carry an excess of 76–190 rare deleterious non-synonymous variants and up to 20 LOF and disease-associated variants. Interestingly, the overall excess of low-frequency variants is similar to that of rare variants (Table 2). Because many variants contributing to disease risk are likely to be segregating at low frequency, we recommend that variant frequency be considered when using the resource to identify pathological candidates.

The combination of variation data with information about regulatory function[13] can potentially improve the power to detect pathological

Table 2 | Per-individual variant load at conserved sites

Variant type	Number of derived variant sites per individual			Excess rare deleterious	Excess low-frequency deleterious
	Derived allele frequency across sample				
	<0.5%	0.5–5%	>5%		
All sites	30–150 K	120–680 K	3.6–3.9 M	ND	ND
Synonymous*	29–120	82–420	1.3–1.4 K	ND	ND
Non-synonymous*	130–400	240–910	2.3–2.7 K	76–190†	77-130†
Stop-gain*	3.9–10	5.3–19	24–28	3.4–7.5†	3.8–11†
Stop-loss	1.0–1.2	1.0–1.9	2.1–2.8	0.81–1.1†	0.80–1.0†
HGMD-DM*	2.5–5.1	4.8–17	11–18	1.6–4.7†	3.8–12†
COSMIC*	1.3–2.0	1.8–5.1	5.2–10	0.93–1.6†	1.3–2.0†
Indel frameshift	1.0–1.3	11–24	60–66	ND§	3.2–11†
Indel non-frameshift	2.1–2.3	9.5–24	67–71	ND§	0–0.73†
Splice site donor	1.7–3.6	2.4–7.2	2.6–5.2	1.6–3.3†	3.1–6.2†
Splice site acceptor	1.5–2.9	1.5–4.0	2.1–4.6	1.4–2.6†	1.2–3.3†
UTR*	120–430	300–1,400	3.5–4.0 K	0–350‡	0–1.2 K‡
Non-coding RNA*	3.9–17	14–70	180–200	0.62–2.6‡	3.4–13‡
Motif gain in TF peak*	4.7–14	23–59	170–180	0–2.6‡	3.8–15‡
Motif loss in TF peak*	18–69	71–300	580–650	7.7–22‡	37–110‡
Other conserved*	2.0–9.9 K	7.1–39 K	120–130 K	ND	ND
Total conserved	2.3–11 K	7.7–42 K	130–150 K	150–510	250–1.3 K

Only sites in which ancestral state can be assigned with high confidence are reported. The ranges reported are across populations. COSMIC, Catalogue of Somatic Mutations in Cancer; HGMD-DM, Human Gene Mutation Database (HGMD) disease-causing mutations; TF, transcription factor; ND, not determined.

* Sites with GERP >2

† Using synonymous sites as a baseline.

‡ Using 'other conserved' as a baseline.

§ Rare indels were filtered in phase I.

non-coding variants. We find that individuals typically contain several thousand variants (and several hundred rare variants) in conserved (GERP conservation score >2) untranslated regions (UTR), non-coding RNAs and transcription-factor-binding motifs (Table 2). Within experimentally defined transcription-factor-binding sites, individuals carry 700–900 conserved motif losses (for the transcription factors analysed, see Supplementary Information), of which 18–69 are rare (<0.5%) and show strong evidence for being selected against. Motif gains are rarer (~200 per individual at conserved sites), but they also show evidence for an excess of rare variants compared with conserved sites with no functional annotation (Table 2). Many of these changes are likely to have weak, slightly deleterious effects on gene regulation and function.

A second major use of the 1000 Genomes Project data in medical genetics is imputing genotypes in existing genome-wide association studies (GWAS)[26]. For common variants, the accuracy of using the phase I data to impute genotypes at sites not on the original GWAS SNP array is typically 90–95% in non-African and approximately 90% in African-ancestry genomes (Fig. 5a and Supplementary Fig. 14a), which is comparable to the accuracy achieved with high-quality benchmark haplotypes (Supplementary Fig. 14b). Imputation accuracy is similar for intergenic SNPs, exome SNPs, indels and large deletions (Supplementary Fig. 14c), despite the different amounts of information about such variants and accuracy of genotypes. For low-frequency variants (1–5%), imputed genotypes have between 60% and 90% accuracy in all populations, including those with admixed ancestry (also comparable to the accuracy from trio-phased haplotypes; Supplementary Fig. 14b).

Imputation has two primary uses: fine-mapping existing association signals and detecting new associations. GWAS have had only a few examples of successful fine-mapping to single causal variants[27,28], often because of extensive haplotype structure within regions of association[29,30]. We find that, in Europeans, each previously reported GWAS signal[31] is, on average, in linkage disequilibrium ($r^2 \geq 0.5$) with 56 variants: 51.5 SNPs and 4.5 indels. In 19% of cases at least one of these variants changes the coding sequence of a nearby gene (compared with 12% in control variants matched for frequency, distance to nearest gene and ascertainment in GWAS arrays) and in 65% of cases at least one of these is at a site with GERP >2 (68% in matched controls). The size of the associated region is typically <200 kb in length (Fig. 5b). Our observations suggest that trans-ethnic fine-mapping experiments are likely to be especially valuable: among the 56 variants that are in strong linkage disequilibrium with a typical GWAS signal, approximately 15 show strong disequilibrium across our four continental groupings (Supplementary Table 15). Our current resource increases the number of variants in linkage disequilibrium with each GWAS signal by 25% compared with the pilot phase of the project and by greater than twofold compared with the HapMap resource.

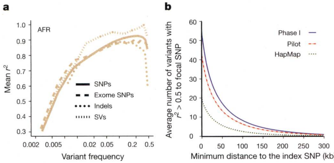


Figure 5 | Implications of phase I 1000 Genomes Project data for GWAS. **a**, Accuracy of imputation of genome-wide SNPs, exome SNPs and indels (using sites on the Illumina 1 M array) into ten individuals of African ancestry (three LWK, four Masaai from Kinyawa, Kenya (MKK), two YRI), sequenced to high coverage by an independent technology[3]. Only indels in regions of high sequence complexity with frequency >1% are analysed. Deletion imputation accuracy estimated by comparison to array data[46] (note that this is for a different set of individuals, although with a similar ancestry, but included on the same plot for clarity). Accuracy measured by squared Pearson correlation coefficient between imputed and true dosage across all sites in a frequency range estimated from the 1000 Genomes data. Lines represent whole-genome SNPs (solid), exome SNPs (long dashes), short indels (dotted) and large deletions (short dashes). SV, structural variants. **b**, The average number of variants in linkage disequilibrium ($r^2 > 0.5$ among EUR) to focal SNPs identified in GWAS[47] as a function of distance from the index SNP. Lines indicate the number of HapMap (green), pilot (red) and phase I (blue) variants.

Discussion

The success of exome sequencing in Mendelian disease genetics[32] and the discovery of rare and low-frequency disease-associated variants in genes associated with complex diseases[27,33,34] strongly support the hypothesis that, in addition to factors such as epistasis[35,36] and gene–environment interactions[37], many other genetic risk factors of substantial effect size remain to be discovered through studies of rare variation. The data generated by the 1000 Genomes Project not only aid the interpretation of all genetic-association studies, but also provide lessons on how best to design and analyse sequencing-based studies of disease.

The use and cost-effectiveness of collecting several data types (low-coverage whole-genome sequence, targeted exome data, SNP genotype data) for finding variants and reconstructing haplotypes are demonstrated here. Exome capture provides private and rare variants that are missed by low-coverage data (approximately 60% of the singleton variants in the sample were detected only from exome data compared with 5% detected only from low-coverage data; Supplementary Fig. 15). However, whole-genome data enable characterization of functional non-coding variation and accurate haplotype estimation, which are essential for the analysis of *cis*-effects around genes, such as those arising from variation in upstream regulatory regions[38]. There are also benefits from integrating SNP array data, for example, to improve genotype estimation[39] and to aid haplotype estimation where array data have been collected on additional family members. In principle, any sources of genotype information (for example, from array CGH) could be integrated using the statistical methods developed here.

Major methodological advances in phase I, including improved methods for detecting and genotyping variants[40], statistical and machine-learning methods for evaluating the quality of candidate variant calls, modelling of genotype likelihoods and performing statistical haplotype integration[41], have generated a high-quality resource. However, regions of low sequence complexity, satellite regions, large repeats and many large-scale structural variants, including copy-number polymorphisms, segmental duplications and inversions (which constitute most of the 'inaccessible genome'), continue to present a major challenge for short-read technologies. Some issues are likely to be improved by methodological developments such as better modelling of read-level errors, integrating *de novo* assembly[42,43] and combining multiple sources of information to aid genotyping of structurally diverse regions[40,44]. Importantly, even subtle differences in data type, data processing or algorithms may lead to systematic differences in false-positive and false-negative error modes between samples. Such differences complicate efforts to compare genotypes between sequencing studies. Moreover, analyses that naively combine variant calls and genotypes across heterogeneous data sets are vulnerable to artefact. Analyses across multiple data sets must therefore either process them in standard ways or use meta-analysis approaches that combine association statistics (but not raw data) across studies.

Finally, the analysis of low-frequency variation demonstrates both the pervasive effects of purifying selection at functionally relevant sites in the genome and how this can interact with population history to lead to substantial local differentiation, even when standard metrics of structure such as F_{ST} are very small. The effect arises primarily because rare variants tend to be recent and thus geographically restricted[6–8]. The implication is that the interpretation of rare variants in individuals with a particular disease should be within the context of the local (either geographic or ancestry-based) genetic background. Moreover, it argues for the value of continuing to sequence individuals from diverse populations to characterize the spectrum of human genetic variation and support disease studies across diverse groups. A further 1,500 individuals from 12 new populations, including at least 15 high-depth trios, will form the final phase of this project.

METHODS SUMMARY

All details concerning sample collection, data generation, processing and analysis can be found in the Supplementary Information. Supplementary Fig. 1 summarizes the process and indicates where relevant details can be found.

Received 4 July; accepted 1 October 2012.

1. Tennessen, J. A. *et al.* Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* **337,** 64–69 (2012).
2. The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467,** 1061–1073 (2010).
3. Drmanac, R. *et al.* Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* **327,** 78–81 (2010).
4. Mills, R. E. *et al.* Mapping copy number variation by population-scale genome sequencing. *Nature* **470,** 59–65 (2011).
5. Marth, G. T. *et al.* The functional spectrum of low-frequency coding variation. *Genome Biol.* **12,** R84 (2011).
6. Nelson, M. R. *et al.* An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science* **337,** 100–104 (2012).
7. Mathieson, I. & McVean, G. Differential confounding of rare and common variants in spatially structured populations. *Nature Genet.* **44,** 243–246 (2012).
8. Gravel, S. *et al.* Demographic history and rare allele sharing among human populations. *Proc. Natl Acad. Sci. USA* **108,** 11983–11988 (2011).
9. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449,** 851–861 (2007).
10. Salmela, E. *et al.* Genome-wide analysis of single nucleotide polymorphisms uncovers population structure in Northern Europe. *PLoS ONE* **3,** e3519 (2008).
11. Lupski, J. R., Belmont, J. W., Boerwinkle, E. & Gibbs, R. A. Clan genomics and the complex architecture of human disease. *Cell* **147,** 32–43 (2011).
12. Lawson, D. J., Hellenthal, G., Myers, S. & Falush, D. Inference of population structure using dense haplotype data. *PLoS Genet.* **8,** e1002453 (2012).
13. ENCODE Project Consortium. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.* **9,** e1001046 (2011).
14. Sasaki, K. *et al.* Expression cloning of a novel Galβ(1–3/1–4)GlcNAc α2,3-sialyltransferase using lectin resistance selection. *J. Biol. Chem.* **268,** 22782–22787 (1993).
15. Marth, G. *et al.* Sequence variations in the public human genome data reflect a bottlenecked population history. *Proc. Natl Acad. Sci. USA* **100,** 376–381 (2003).
16. Keinan, A. & Clark, A. G. Recent explosive human population growth has resulted in an excess of rare genetic variants. *Science* **336,** 740–743 (2012).
17. Hall, G. M. *Slavery and African Ethnicities in the Americas: Restoring the Links* (Univ. North Carolina Press, 2005).
18. Bryc, K. *et al.* Genome-wide patterns of population structure and admixture in West Africans and African Americans. *Proc. Natl Acad. Sci. USA* **107,** 786–791 (2010).
19. Davydov, E. V. *et al.* Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLOS Comput. Biol.* **6,** e1001025 (2010).
20. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40,** D109–D114 (2012).
21. Kim, T. H. *et al.* Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* **128,** 1231–1245 (2007).
22. Bamshad, M. J. *et al.* Exome sequencing as a tool for Mendelian disease gene discovery. *Nature Rev. Genet.* **12,** 745–755 (2011).
23. Cancer Genome Altas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **474,** 609–615 (2011).
24. Stenson, P. D. *et al.* The Human Gene Mutation Database: 2008 update. *Genome Med.* **1,** 13 (2009).
25. Forbes, S. A. *et al.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39,** D945–D950 (2011).
26. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* **1,** 457–470 (2011).
27. Sanna, S. *et al.* Fine mapping of five loci associated with low-density lipoprotein cholesterol detects variants that double the explained heritability. *PLoS Genet.* **7,** e1002198 (2011).
28. Gregory, A. P., Dendrou, C. A., Bell, J., McVean, G. & Fugger, L. TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. *Nature* **488,** 508–511 (2012).
29. Hassanein, M. T. *et al.* Fine mapping of the association with obesity at the *FTO* locus in African-derived populations. *Hum. Mol. Genet.* **19,** 2907–2916 (2010).
30. Maller, J., The Wellcome Trust Case Control Consortium. Fine mapping of 14 loci identified through genome-wide association analyses. *Nature Genet.* (in the press).
31. Hindorff, L. A. *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl Acad. Sci. USA* **106,** 9362–9367 (2009).
32. Bamshad, M. J. *et al.* The Centers for Mendelian Genomics: A new large-scale initiative to identify the genes underlying rare Mendelian conditions. *Am. J. Med. Genet. A.* (2012).
33. Momozawa, Y. *et al.* Resequencing of positional candidates identifies low frequency *IL23R* coding variants protecting against inflammatory bowel disease. *Nature Genet.* **43,** 43–47 (2011).
34. Raychaudhuri, S. *et al.* A rare penetrant mutation in *CFH* confers high risk of age-related macular degeneration. *Nature Genet.* **43,** 1232–1236 (2011).
35. Strange, A. *et al.* A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between *HLA-C* and *ERAP1*. *Nature Genet.* **42,** 985–990 (2010).
36. Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc. Natl Acad. Sci. USA* **109,** 1193–1198 (2012).
37. Thomas, D. Gene-environment-wide association studies: emerging approaches. *Nature Rev. Genet.* **11,** 259–272 (2010).
38. Degner, J. F. *et al.* DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature* **482,** 390–394 (2012).
39. Flannick, J. *et al.* Efficiency and power as a function of sequence coverage, SNP array density, and imputation. *PLOS Comput. Biol.* **8,** e1002604 (2012).
40. Handsaker, R. E., Korn, J. M., Nemesh, J. & McCarroll, S. A. Discovery and genotyping of genome structural polymorphism by sequencing on a population scale. *Nature Genet.* **43,** 269–276 (2011).
41. Li, Y., Sidore, C., Kang, H. M., Boehnke, M. & Abecasis, G. R. Low-coverage sequencing: implications for design of complex trait association studies. *Genome Res.* **21,** 940–951 (2011).
42. Iqbal, Z., Caccamo, M., Turner, I., Flicek, P. & McVean, G. *De novo* assembly and genotyping of variants using colored de Bruijn graphs. *Nature Genet.* **44,** 226–232 (2012).
43. Simpson, J. T. & Durbin, R. Efficient construction of an assembly string graph using the FM-index. *Bioinformatics* **26,** i367–i373 (2010).
44. Sudmant, P. H. *et al.* Diversity of human copy number variation and multicopy genes. *Science* **330,** 641–646 (2010).
45. Chambers, J. C. *et al.* Genetic loci influencing kidney function and chronic kidney disease. *Nature Genet.* **42,** 373–375 (2010).
46. Conrad, D. F. *et al.* Origins and functional impact of copy number variation in the human genome. *Nature* **464,** 704–712 (2010).
47. Hindorff, L. A. *et al. A Catalog of Published Genome-Wide Association Studies.* Available at http://www.genome.gov/gwastudies (accessed, September 2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank many people who contributed to this project: A. Naranjo, M. V. Parra and C. Duque for help with the collection of the Colombian samples; N. Kälin and F. Laplace for discussions; A. Schlattl and T. Zichner for assistance in managing data sets; E. Appelbaum, H. Arbery, E. Birney, S. Bumpstead, J. Camarata, J. Carey, G. Cochrane, M. DaSilva, S. Dökel, E. Drury, C. Duque, K. Gyaltsen, P. Jokinen, B. Lenz, S. Lewis, D. Lu, A. Naranjo, S. Ott, I. Padioleau, M. V. Parra, N. Patterson, A. Price, L. Sadzewicz, S. Schrinner, N. Sengamalay, J. Sullivan, F. Ta, Y. Vaydylevich, O. Venn, K. Watkins and A. Yurovsky for assistance, discussion and advice. We thank the people who generously contributed their samples, from these populations: Yoruba in Ibadan, Nigeria; the Han Chinese in Beijing, China; the Japanese in Tokyo, Japan; the Utah CEPH community; the Luhya in Webuye, Kenya; people with African ancestry in the Southwest United States; the Toscani in Italia; people with Mexican ancestry in Los Angeles, California; the Southern Han Chinese in China; the British in England and Scotland; the Finnish in Finland; the Iberian Populations in Spain; the Colombians in Medellin, Colombia; and the Puerto Ricans in Puerto Rico. This research was supported in part by Wellcome Trust grants WT098051 to R.M.D., M.E.H. and C.T.S.; WT090532/Z/09/Z, WT085475/Z/08/Z and WT095552/Z/11/Z to P.Do.; WT086084/Z/08/Z and WT090532/Z/09/Z to G.A.M.; WT089250/Z/09/Z to I.M.; WT085532AIA to P.F.; Medical Research Council grant G0900747(91070) to G.A.M.; British Heart Foundation grant RG/09/12/28096 to C.A.A.; the National Basic Research Program of China (973 program no. 2011CB809201, 2011CB809202 and 2011CB809203); the Chinese 863 program (2012AA02A201); the National Natural Science Foundation of China (30890032, 31161130357); the Shenzhen Key Laboratory of Transomics Biotechnologies (CXB201108250096A); the Shenzhen Municipal Government of China (grants ZYC200903240080A and ZYC201105170397A); Guangdong Innovative Research Team Program (no. 2009010016); BMBF grant 01GS08201 to H.Le.; BMBF grant 0315428A to R.H.; the Max Planck Society; Swiss National Science Foundation 31003A_130342 to E.T.D.; Swiss National Science Foundation NCCR 'Frontiers in Genetics' grant to E.T.D.; Louis Jeantet Foundation grant to E.T.D.; Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/I021213/1 to A.R.-L.; German Research Foundation (Emmy Noether Fellowship KO 4037/1-1) to J.O.K.; Netherlands Organization for Scientific Research VENI grant 639.021.125 to K.Y.; Beatriu de Pinos Program grants 2006BP-A 10144 and 2009BP-B 00274 to M.V.; Israeli Science Foundation grant 04514831 to E.H.; Genome Québec and the Ministry of Economic Development, Innovation and Trade grant PSR-SIIRI-195 to P.Aw.; National Institutes of Health (NIH) grants UO1HG5214, RC2HG5581 and RO1MH84698 to G.R.A.; R01HG4719 and R01HG3698 to G.T.M; RC2HG5552 and UO1HG6513 to G.R.A. and G.T.M.; R01HG4960 and R01HG5701 to B.L.B.; U01HG5715 to C.D.B. and A.G.C.; T32GM8283 to D.Cl.; U01HG5208 to M.J.D.; U01HG6569 to M.A.D.; R01HG2898 and R01CA166661 to S.E.D.; UO1HG5209, UO1HG5725 and P41HG4221 to C.Le.; P01HG4120 to E.E.E.; U01HG5728 to Yu.F.; U54HG3273 and U01HG5211 to R.A.G.;

R01HL95045 to S.B.G.; U41HG4568 to S.J.K.; P41HG2371 to W.J.K.; ES015794, AI077439, HL088133 and HL078885 to E.G.B.; RC2HL102925 to S.B.G. and D.M.A.; R01GM59290 to L.B.J. and M.A.B.; U54HG3067 to E.S.L. and S.B.G.; T15LM7033 to B.K.M.; T32HL94284 to J.L.R.-F.; DP2OD6514 and BAA-NIAID-DAIT-NIHAI2009061 to P.C.S.; T32GM7748 to X.S.; U54HG3079 to R.K.W.; UL1RR024131 to R.D.H.; HHSN268201100040C to the Coriell Institute for Medical Research; a Sandler Foundation award and an American Asthma Foundation award to E.G.B.; an IBM Open Collaborative Research Program award to Y.B.; an A.G. Leventis Foundation scholarship to D.K.X.; a Wolfson Royal Society Merit Award to P.Do.; a Howard Hughes Medical Institute International Fellowship award to P.H.S.; a grant from T. and V. Stanley to S.C.Y.; and a Mary Beryl Patch Turnbull Scholar Program award to K.C.B. E.H. is a faculty fellow of the Edmond J. Safra Bioinformatics program at Tel-Aviv University. E.E.E. and D.H. are investigators of the Howard Hughes Medical Institute. M.V.G. is a long-term fellow of EMBO.

Author Contributions Details of author contributions can be found in the author list.

Author Information All primary data, alignments, individual call sets, consensus call sets, integrated haplotypes with genotype likelihoods and supporting data including details of validation are available from the project website (http://www.1000genomes.org). Variant and haplotypes for specific genomic regions and specific samples can be viewed and downloaded through the project browser (http://browser.1000genomes.org/). Common project variants with no known medical impact have been compiled by dbSNP for filtering (http://www.ncbi.nlm.nih.gov/variation/docs/human_variation_vcf/). The authors declare competing financial interests: details are available in the online version of the paper. Reprints and permissions information is available at www.nature.com/reprints. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.A.M. (mcvean@well.ox.ac.uk). This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported licence. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-sa/3.0/

The 1000 Genomes Consortium (Participants are arranged by project role, then by institution alphabetically, and finally alphabetically within institutions except for Principal Investigators and Project Leaders, as indicated.)

Corresponding author Gil A. McVean[1,2]

Steering committee David M. Altshuler[3,4,5] (Co-Chair), Richard M. Durbin[6] (Co-Chair), Gonçalo R. Abecasis[7], David R. Bentley[8], Aravinda Chakravarti[9], Andrew G. Clark[10], Peter Donnelly[1,2], Evan E. Eichler[11], Paul Flicek[12], Stacey B. Gabriel[3], Richard A. Gibbs[13], Eric D. Green[14], Matthew E. Hurles[6], Bartha M. Knoppers[15], Jan O. Korbel[16], Eric S. Lander[3], Charles Lee[17], Hans Lehrach[18,19], Elaine R. Mardis[20], Gabor T. Marth[21], Gil A. McVean[1,2], Deborah A. Nickerson[22], Jeanette P. Schmidt[23], Stephen T. Sherry[24], Jun Wang[25,26,27], Richard K. Wilson[20]

Production group: Baylor College of Medicine Richard A. Gibbs[13] (Principal Investigator), Huyen Dinh[13], Christie Kovar[13], Sandra Lee[13], Lora Lewis[13], Donna Muzny[13], Jeff Reid[13], Min Wang[13]; **BGI-Shenzhen** Jun Wang[25,26,27] (Principal Investigator), Xiaodong Fang[25], Xiaosen Guo[25], Min Jian[25], Hui Jiang[25], Xin Jin[25], Guoqing Li[25], Jingxiang Li[25], Yingrui Li[25], Zhuo Li[25], Xiao Liu[25], Yao Lu[25], Xuedi Ma[25], Zhe Su[25], Shuaishuai Tai[25], Meifang Tang[25], Bo Wang[25], Guangbiao Wang[25], Honglong Wu[25], Renhua Wu[25], Ye Yin[25], Wenwei Zhang[25], Jiao Zhao[25], Meiru Zhao[25], Xiaole Zheng[25], Yan Zhou[25]; **Broad Institute of MIT and Harvard** Eric S. Lander[3] (Principal Investigator), David M. Altshuler[3,4,5], Stacey B. Gabriel[3] (Co-Chair), Namrata Gupta[3]; **European Bioinformatics Institute** Paul Flicek[12] (Principal Investigator), Laura Clarke[12], Rasko Leinonen[12], Richard E. Smith[12], Xiangqun Zheng-Bradley[12]; **Illumina** David R. Bentley[8] (Principal Investigator), Russell Grocock[8], Sean Humphray[8], Terena James[8], Zoya Kingsbury[8]; **Max Planck Institute for Molecular Genetics** Hans Lehrach[18,19] (Principal Investigator), Ralf Sudbrak[18] (Project Leader), Marcus W. Albrecht[28], Vyacheslav S. Amstislavskiy[18], Tatiana A. Borodina[28], Matthias Lienhard[18], Florian Mertes[18], Marc Sultan[18], Bernd Timmermann[18], Marie-Laure Yaspo[18]; **US National Institutes of Health** Stephen T. Sherry[24] (Principal Investigator); **University of Oxford** Gil A. McVean[1,2] (Principal Investigator); **Washington University in St Louis** Elaine R. Mardis[20] (Co-Principal Investigator) (Co-Chair), Richard K. Wilson[20] (Co-Principal Investigator), Lucinda Fulton[20], Robert Fulton[20], George M. Weinstock[20]; **Wellcome Trust Sanger Institute** Richard M. Durbin[6] (Principal Investigator), Senduran Balasubramaniam[6], John Burton[6], Petr Danecek[6], Thomas M. Keane[6], Anja Kolb-Kokocinski[6], Shane McCarthy[6], James Stalker[6], Michael Quail[6]

Analysis group: Affymetrix Jeanette P. Schmidt[23] (Principal Investigator), Christopher J. Davies[23], Jeremy Gollub[23], Teresa Webster[23], Brant Wong[23], Yiping Zhan[23]; **Albert Einstein College of Medicine** Adam Auton[29] (Principal Investigator); **Baylor College of Medicine** Richard A. Gibbs[13] (Principal Investigator), Fuli Yu[13] (Project Leader), Matthew Bainbridge[13], Danny Challis[13], Uday S. Evani[13], James Lu[13], Donna Muzny[13], Uma Nagaswamy[13], Jeff Reid[13], Aniko Sabo[13], Yi Wang[13], Jin Yu[13]; **BGI-Shenzhen** Jun Wang[25,26,27] (Principal Investigator), Lachlan J. M. Coin[25], Lin Fang[25], Xiaosen Guo[25], Xin Jin[25], Guoqing Li[25], Qibin Li[25], Yingrui Li[25], Zhenyu Li[25], Haoxiang Lin[25], Binghang Liu[25], Ruibang Luo[25], Nan Qin[25], Haojing Shao[25], Bingqiang Wang[25], Yinlong Xie[25], Chen Ye[25], Chang Yu[25], Fan Zhang[25], Hancheng Zheng[25], Hongmei Zhu[25]; **Boston College** Gabor T. Marth[21] (Principal Investigator), Erik P. Garrison[21], Deniz Kural[21], Wan-Ping Lee[21], Wen Fung Leong[21], Alistair N. Ward[21], Jiantao Wu[21], Mengyao Zhang[21]; **Brigham and Women's Hospital** Charles Lee[17] (Principal Investigator), Lauren Griffin[17], Chih-Heng Hsieh[17], Ryan E. Mills[17,30], Xinghua Shi[17], Marcin von Grotthuss[17], Chengsheng Zhang[17]; **Broad Institute of MIT and Harvard** Mark J. Daly[3] (Principal Investigator), Mark A. DePristo[3] (Project Leader), David M. Altshuler[3,4,5], Eric Banks[3], Gaurav Bhatia[3], Mauricio O. Carneiro[3], Guillermo del Angel[3], Stacey B. Gabriel[3], Giulio Genovese[3], Namrata Gupta[3], Robert E. Handsaker[3,5], Chris Hartl[3], Eric S. Lander[3], Steven A. McCarroll[3], James C. Nemesh[3], Ryan E. Poplin[3], Stephen F. Schaffner[3], Khalid Shakir[3]; **Cold Spring Harbor Laboratory** Seungtai C. Yoon[31] (Principal Investigator), Jayon Lihm[31], Vladimir Makarov[32]; **Dankook University** Hanjun Jin[33] (Principal Investigator), Wook Kim[34], Ki Cheol Kim[34]; **European Molecular Biology Laboratory** Jan O. Korbel[16] (Principal Investigator), Tobias Rausch[16]; **European Bioinformatics Institute** Paul Flicek[12] (Principal Investigator), Kathryn Beal[12], Laura Clarke[12], Fiona Cunningham[12], Javier Herrero[12], William M. McLaren[12], Graham R. S. Ritchie[12], Richard E. Smith[12], Xiangqun Zheng-Bradley[12]; **Cornell University** Andrew G. Clark[10] (Principal Investigator), Srikanth Gottipati[35], Alon Keinan[10], Juan L. Rodriguez-Flores[10]; **Harvard University** Pardis C. Sabeti[3,36] (Principal Investigator), Sharon R. Grossman[3,36], Shervin Tabrizi[3,36], Ridhi Tariyal[3,36]; **Human Gene Mutation Database** David N. Cooper[37] (Principal Investigator), Edward V. Ball[37], Peter D. Stenson[37]; **Illumina** David R. Bentley[8] (Principal Investigator), Bret Barnes[38], Markus Bauer[8], R. Keira Cheetham[8], Tony Cox[8], Michael Eberle[8], Sean Humphray[8], Scott Kahn[38], Lisa Murray[8], John Peden[8], Richard Shaw[8]; **Leiden University Medical Center** Kai Ye[39] (Principal Investigator); **Louisiana State University** Mark A. Batzer[40] (Principal Investigator), Miriam K. Konkel[40], Jerilyn A. Walker[40]; **Massachusetts General Hospital** Daniel G. MacArthur[41] (Principal Investigator), Monkol Lek[41]; **Max Planck Institute for Molecular Genetics** Ralf Sudbrak[18] (Project Leader), Vyacheslav S. Amstislavskiy[18], Ralf Herwig[18]; **Pennsylvania State University** Mark D. Shriver[42] (Principal Investigator); **Stanford University** Carlos D. Bustamante[43] (Principal Investigator), Jake K. Byrnes[44], Francisco M. De La Vega[10], Simon Gravel[43], Eimear E. Kenny[43], Jeffrey M. Kidd[43], Phil Lacroute[43], Brian K. Maples[43], Andres Moreno-Estrada[43], Fouad Zakharia[43]; **Tel-Aviv University** Eran Halperin[45,46,47] (Principal Investigator), Yael Baran[45]; **Translational Genomics Research Institute** David W. Craig[48] (Principal Investigator), Alexis Christoforides[48], Nils Homer[49], Tyler Izatt[48], Ahmet A. Kurdoglu[48], Shripad A. Sinari[48], Kevin Squire[50]; **US National Institutes of Health** Stephen T. Sherry[24] (Principal Investigator), Chunlin Xiao[24]; **University of California, San Diego** Jonathan Sebat[51,52] (Principal Investigator), Vineet Bafna[53], Kenny Ye[54]; **University of California, San Francisco** Esteban G. Burchard[55] (Principal Investigator), Ryan D. Hernandez[55] (Principal Investigator), Christopher R. Gignoux[55]; **University of California, Santa Cruz** David Haussler[56,57] (Principal Investigator), Sol J. Katzman[56], W. James Kent[56]; **University of Chicago** Bryan Howie[58]; **University College London** Andres Ruiz-Linares[59] (Principal Investigator); **University of Geneva** Emmanouil T. Dermitzakis[60,61,62] (Principal Investigator), Tuuli Lappalainen[60,61,62]; **University of Maryland School of Medicine** Scott E. Devine[63] (Principal Investigator), Xinyue Liu[63], Ankit Maroo[63], Luke J. Tallon[63]; **University of Medicine and Dentistry of New Jersey** Jeffrey A. Rosenfeld[64,65] (Principal Investigator), Leslie P. Michelson[64]; **University of Michigan** Gonçalo R. Abecasis[7] (Principal Investigator) (Co-Chair), Hyun Min Kang[7] (Project Leader), Paul Anderson[7], Andrea Angius[66], Abigail Bigham[67], Tom Blackwell[7], Fabio Busonero[7,66,68], Francesco Cucca[66,68], Christian Fuchsberger[7], Chris Jones[69], Goo Jun[7], Yun Li[70], Robert Lyons[71], Andrea Maschio[7,66,68], Eleonora Porcu[7,66,68], Fred Reinier[69], Serena Sanna[66], David Schlessinger[72], Carlo Sidore[7,66,68], Adrian Tan[7], Mary Kate Trost[7]; **University of Montréal** Philip Awadalla[73] (Principal Investigator), Alan Hodgkinson[73]; **University of Oxford** Gerton Lunter[1] (Principal Investigator), Gil A. McVean[1,2] (Principal Investigator) (Co-Chair), Jonathan L. Marchini[1,2] (Principal Investigator), Simon Myers[1,2] (Principal Investigator), Claire Churchhouse[2], Olivier Delaneau[2], Anjali Gupta-Hinch[1], Zamin Iqbal[1], Iain Mathieson[1], Andy Rimmer[1], Dionysia K. Xifara[1,2]; **University of Puerto Rico** Taras K. Oleksyk[74] (Principal Investigator); **University of Texas Health Sciences Center at Houston** Yunxin Fu[75] (Principal Investigator), Xiaoming Liu[75], Momiao Xiong[75]; **University of Utah** Lynn Jorde[76] (Principal Investigator), David Witherspoon[76], Jinchuan Xing[77]; **University of Washington** Evan E. Eichler[11] (Principal Investigator), Brian L. Browning[78] (Principal Investigator), Can Alkan[22,79], Iman Hajirasouliha[80], Fereydoun Hormozdiari[22], Arthur Ko[22], Peter H. Sudmant[22]; **Washington University in St Louis** Elaine R. Mardis[20] (Co-Principal Investigator), Ken Chen[81], Asif Chinwalla[20], Li Ding[20], David Dooling[20], Daniel C. Koboldt[20], Michael D. McLellan[20], John W. Wallis[20], Michael C. Wendl[20], Qunyuan Zhang[20]; **Wellcome Trust Sanger Institute** Richard M. Durbin[6] (Principal Investigator), Matthew E. Hurles[6] (Principal Investigator), Chris Tyler-Smith[6] (Principal Investigator), Cornelis A. Albers[82], Qasim Ayub[6], Senduran Balasubramaniam[6], Yuan Chen[6], Alison J. Coffey[6], Vincenza Colonna[6,83], Petr Danecek[6], Ni Huang[6], Luke Jostins[6], Thomas M. Keane[6], Heng Li[3,6], Shane McCarthy[6], Aylwyn Scally[6], James Stalker[6], Klaudia Walter[6], Yali Xue[6], Yujun Zhang[6]; **Yale University** Mark B. Gerstein[84,85,86] (Principal Investigator), Alexej Abyzov[84,86], Suganthi Balasubramanian[86], Jieming Chen[84], Declan Clarke[87], Yao Fu[84], Lukas Habegger[84], Arif O. Harmanci[84], Mike Jin[86], Ekta Khurana[86], Xinmeng Jasmine Mu[84], Cristina Sisu[84]

Structural variation group: BGI-Shenzhen Yingrui Li[25], Ruibang Luo[25], Hongmei Zhu[25]; **Brigham and Women's Hospital** Charles Lee[17] (Principal Investigator) (Co-Chair), Lauren Griffin[17], Chih-Heng Hsieh[17], Ryan E. Mills[17,30], Xinghua Shi[17], Marcin von Grotthuss[17], Chengsheng Zhang[17]; **Boston College** Gabor T. Marth[21] (Principal Investigator), Erik P. Garrison[21], Deniz Kural[21], Wan-Ping Lee[21], Alistair N. Ward[21], Jiantao Wu[21], Mengyao Zhang[21]; **Broad Institute of MIT and Harvard** Steven A. McCarroll[3] (Project Leader), David M. Altshuler[3,4,5], Eric Banks[3], Guillermo del Angel[3], Giulio Genovese[3], Robert E. Handsaker[3,5], Chris Hartl[3], James C. Nemesh[3], Khalid Shakir[3]; **Cold Spring Harbor Laboratory** Seungtai C. Yoon[31] (Principal Investigator), Jayon Lihm[31], Vladimir Makarov[32]; **Cornell University** Jeremiah Degenhardt[10]; **European Bioinformatics Institute** Paul Flicek[12] (Principal Investigator), Laura Clarke[12], Richard E. Smith[12], Xiangqun Zheng-Bradley[12];

European Molecular Biology Laboratory Jan O. Korbel[16] (Principal Investigator) (Co-Chair), Tobias Rausch[16], Adrian M. Stütz[16]; **Illumina** David R. Bentley[8] (Principal Investigator), Bret Barnes[38], R. Keira Cheetham[8], Michael Eberle[8], Sean Humphray[8], Scott Kahn[38], Lisa Murray[8], Richard Shaw[8]; **Leiden University Medical Center** Kai Ye[39] (Principal Investigator); **Louisiana State University** Mark A. Batzer[40] (Principal Investigator), Miriam K. Konkel[40], Jerilyn A. Walker[40]; **Stanford University** Phil Lacroute[43]; **Translational Genomics Research Institute** David W. Craig[48] (Principal Investigator), Nils Homer[49]; **US National Institutes of Health** Deanna Church[24], Chunlin Xiao[24]; **University of California, San Diego** Jonathan Sebat[51,52] (Principal Investigator), Vineet Bafna[53], Jacob J. Michaelson[88], Kenny Ye[54]; **University of Maryland School of Medicine** Scott E. Devine[63] (Principal Investigator), Xinyue Liu[63], Ankit Maroo[63], Luke J. Tallon[63]; **University of Oxford** Gerton Lunter[1] (Principal Investigator), Gil A. McVean[1,2] (Principal Investigator), Zamin Iqbal[1]; **University of Utah** David Witherspoon[76], Jinchuan Xing[77]; **University of Washington** Evan E. Eichler[11] (Principal Investigator) (Co-Chair), Can Alkan[22,79], Iman Hajirasouliha[80], Fereydoun Hormozdiari[22], Arthur Ko[22], Peter H. Sudmant[22]; **Washington University in St Louis** Ken Chen[81], Asif Chinwalla[20], Li Ding[20], Michael D. McLellan[20], John W. Wallis[20]; **Wellcome Trust Sanger Institute** Matthew E. Hurles[6] (Principal Investigator) (Co-Chair), Ben Blackburne[6], Heng Li[3,6], Sarah J. Lindsay[6], Zemin Ning[6], Aylwyn Scally[6], Klaudia Walter[6], Yujun Zhang[6]; **Yale University** Mark B. Gerstein[84,85,86] (Principal Investigator), Alexej Abyzov[84,86], Jieming Chen[84], Declan Clarke[87], Ekta Khurana[86], Xinmeng Jasmine Mu[84], Cristina Sisu[84]

Exome group: Baylor College of Medicine Richard A. Gibbs[13] (Principal Investigator) (Co-Chair), Fuli Yu[13] (Project Leader), Matthew Bainbridge[13], Danny Challis[13], Uday S. Evani[13], Christie Kovar[13], Lora Lewis[13], James Lu[13], Donna Muzny[13], Uma Nagaswamy[13], Jeff Reid[13], Aniko Sabo[13], Jin Yu[13]; **BGI-Shenzhen** Xiaosen Guo[25], Yingrui Li[25], Renhua Wu[25]; **Boston College** Gabor T. Marth[21] (Principal Investigator) (Co-Chair), Erik P. Garrison[21], Wen Fung Leong[21], Alistair N. Ward[21]; **Broad Institute of MIT and Harvard** Guillermo del Angel[3], Mark A. DePristo[3], Stacey B. Gabriel[3], Namrata Gupta[3], Chris Hartl[3], Ryan E. Poplin[3]; **Cornell University** Andrew G. Clark[10] (Principal Investigator), Juan L. Rodriguez-Flores[10]; **European Bioinformatics Institute** Paul Flicek[12] (Principal Investigator), Laura Clarke[12], Richard E. Smith[12], Xiangqun Zheng-Bradley[12]; **Massachusetts General Hospital** Daniel G. MacArthur[41] (Principal Investigator); **Stanford University** Carlos D. Bustamante[43] (Principal Investigator), Simon Gravel[43]; **Translational Genomics Research Institute** David W. Craig[48] (Principal Investigator), Alexis Christoforides[48], Nils Homer[49], Tyler Izatt[48]; **US National Institutes of Health** Stephen T. Sherry[24] (Principal Investigator), Chunlin Xiao[24]; **University of Geneva** Emmanouil T. Dermitzakis[60,61,62] (Principal Investigator); **University of Michigan** Gonçalo R. Abecasis[7] (Principal Investigator), Hyun Min Kang[7]; **University of Oxford** Gil A. McVean[1,2] (Principal Investigator); **Washington University in St Louis** Elaine R. Mardis[20] (Principal Investigator), David Dooling[20], Lucinda Fulton[20], Robert Fulton[20], Daniel C. Koboldt[20]; **Wellcome Trust Sanger Institute** Richard M. Durbin[6] (Principal Investigator), Senduran Balasubramaniam[6], Thomas M. Keane[6], Shane McCarthy[6], James Stalker[6]; **Yale University** Mark B. Gerstein[84,85,86] (Principal Investigator), Suganthi Balasubramanian[86], Lukas Habegger[84]

Functional interpretation group: Boston College Erik P. Garrison[21]; **Baylor College of Medicine** Richard A. Gibbs[13] (Principal Investigator), Matthew Bainbridge[13], Donna Muzny[13], Fuli Yu[13], Jin Yu[13]; **Broad Institute of MIT and Harvard** Guillermo del Angel[3], Robert E. Handsaker[3,5]; **Cold Spring Harbor Laboratory** Vladimir Makarov[32]; **Cornell University** Juan L. Rodriguez-Flores[10]; **Dankook University** Hanjun Jin[33] (Principal Investigator), Wook Kim[34], Ki Cheol Kim[34]; **European Bioinformatics Institute** Paul Flicek[12] (Principal Investigator), Kathryn Beal[12], Laura Clarke[12], Fiona Cunningham[12], Javier Herrero[12], William M. McLaren[12], Graham R. S. Ritchie[12], Xiangqun Zheng-Bradley[12]; **Harvard University** Shervin Tabrizi[3,36], **Massachusetts General Hospital** Daniel G. MacArthur[41] (Principal Investigator), Monkol Lek[41]; **Stanford University** Carlos D. Bustamante[43] (Principal Investigator), Francisco M. De La Vega[10]; **Translational Genomics Research Institute** David W. Craig[48] (Principal Investigator), Ahmet A. Kurdoglu[48]; **University of Geneva** Tuuli Lappalainen[60,61,62]; **University of Medicine and Dentistry of New Jersey** Jeffrey A. Rosenfeld[64,65] (Principal Investigator), Leslie P. Michelson[64,65]; **University of Montréal** Philip Awadalla[73] (Principal Investigator), Alan Hodgkinson[73]; **University of Oxford** Gil A. McVean[1,2] (Principal Investigator); **Washington University in St Louis** Ken Chen[81]; **Wellcome Trust Sanger Institute** Chris Tyler-Smith[6] (Principal Investigator) (Co-Chair), Yuan Chen[6], Vincenza Colonna[6,83], Adam Frankish[6], Jennifer Harrow[6], Yali Xue[6]; **Yale University** Mark B. Gerstein[84,85,86] (Principal Investigator) (Co-Chair), Alexej Abyzov[84,86], Suganthi Balasubramanian[86], Jieming Chen[84], Declan Clarke[87], Yao Fu[84], Arif O. Harmanci[84], Mike Jin[86], Ekta Khurana[86], Xinmeng Jasmine Mu[84], Cristina Sisu[84]

Data coordination centre group: Baylor College of Medicine Richard A. Gibbs[13] (Principal Investigator), Gerald Fowler[13], Walker Hale[13], Divya Kalra[13], Christie Kovar[13], Donna Muzny[13], Jeff Reid[13]; **BGI-Shenzhen** Jun Wang[25,26,27] (Principal Investigator), Xiaosen Guo[25], Guoqing Li[25], Yingrui Li[25], Xiaole Zheng[25]; **Broad Institute of MIT and Harvard** David M. Altshuler[3,4,5]; **European Bioinformatics Institute** Paul Flicek[12] (Principal Investigator) (Co-Chair), Laura Clarke[12] (Project Leader), Jonathan Barker[12], Gavin Kelman[12], Eugene Kulesha[12], Rasko Leinonen[12], William M. McLaren[12], Rajesh Radhakrishnan[12], Asier Roa[12], Dmitriy Smirnov[12], Richard E. Smith[12], Ian Streeter[12], Iliana Toneva[12], Brendan Vaughan[12], Xiangqun Zheng-Bradley[12]; **Illumina** David R. Bentley[8] (Principal Investigator), Tony Cox[8], Sean Humphray[8], Scott Kahn[38]; **Max Planck Institute for Molecular Genetics** Ralf Sudbrak[18] (Project Leader), Marcus W. Albrecht[28], Matthias Lienhard[18]; **Translational Genomics Research Institute** David W. Craig[48] (Principal Investigator), Tyler Izatt[48], Ahmet A. Kurdoglu[48]; **US National Institutes of Health** Stephen T. Sherry[24] (Principal Investigator) (Co-Chair), Victor Ananiev[24], Zinaida Belaia[24], Dimitriy Beloslyudtsev[24], Nathan Bouk[24], Chao Chen[24], Deanna Church[24], Robert Cohen[24], Charles Cook[24], John Garner[24], Timothy Hefferon[24], Mikhail Kimelman[24], Chunlei Liu[24], John Lopez[24], Peter Meric[24], Chris O'Sullivan[89], Yuri Ostapchuk[24], Lon Phan[24], Sergiy Ponomarov[24], Valerie Schneider[24], Eugene Shekhtman[24], Karl Sirotkin[24], Douglas Slotta[24], Chunlin Xiao[24], Hua Zhang[24]; **University of California, Santa Cruz** David Haussler[56,57] (Principal Investigator); **University of Michigan** Gonçalo R. Abecasis[7] (Principal Investigator); **University of Oxford** Gil A. McVean[1,2] (Principal Investigator); **University of Washington** Can Alkan[22,79], Arthur Ko[22]; **Washington University in St Louis** David Dooling[20]; **Wellcome Trust Sanger Institute** Richard M. Durbin[6] (Principal Investigator), Senduran Balasubramaniam[6], Thomas M. Keane[6], Shane McCarthy[6], James Stalker[6]

Samples and ELSI group Aravinda Chakravarti[9] (Co-Chair), Bartha M. Knoppers[15] (Co-Chair), Gonçalo R. Abecasis[7], Kathleen C. Barnes[90], Christine Beiswanger[91], Esteban G. Burchard[55], Carlos D. Bustamante[43], Hongyu Cai[25], Hongzhi Cao[25], Richard M. Durbin[6], Neda Gharani[91], Richard A. Gibbs[13], Christopher R. Gignoux[55], Simon Gravel[43], Brenna Henn[43], Danielle Jones[35], Lynn Jorde[76], Jane S. Kaye[92], Alon Keinan[10], Alastair Kent[93], Angeliki Kerasidou[1], Yingrui Li[25], Rasika Mathias[94], Gil A. McVean[1,2], Andres Moreno-Estrada[43], Pilar N. Ossorio[95,96], Michael Parker[97], David Reich[5], Charles N. Rotimi[98], Charmaine D. Royal[99], Karla Sandoval[43], Yeyang Su[25], Ralf Sudbrak[18], Zhongming Tian[25], Bernd Timmermann[18], Sarah Tishkoff[100], Lorraine H. Toji[91], Chris Tyler-Smith[6], Marc Via[101], Yuhong Wang[25], Huanming Yang[25], Ling Yang[25], Jiayong Zhu[25]

Sample collection: British from England and Scotland (GBR) Walter Bodmer[102]; **Colombians in Medellín, Colombia (CLM)** Gabriel Bedoya[103], Andres Ruiz-Linares[59]; **Han Chinese South (CHS)** Cai Zhi Ming[25], Gao Yang[104], Chu Jia You[105]; **Finnish in Finland (FIN)** Leena Peltonen‡; **Iberian populations in Spain (IBS)** Andres Garcia-Montero[106], Alberto Orfao[107]; **Puerto Ricans in Puerto Rico (PUR)** Julie Dutil[108], Juan C. Martinez-Cruzado[74], Taras K. Oleksyk[74]

Scientific management Lisa D. Brooks[109], Adam L. Felsenfeld[109], Jean E. McEwen[109], Nicholas C. Clemm[109], Audrey Duncanson[110], Michael Dunn[110], Eric D. Green[14], Mark S. Guyer[109], Jane L. Peterson[109]

Writing group Goncalo R. Abecasis[7], Adam Auton[29], Lisa D. Brooks[109], Mark A. DePristo[3], Richard M. Durbin[6], Robert E. Handsaker[3,5], Hyun Min Kang[7], Gabor T. Marth[21], Gil A. McVean[1,2]

[1]Wellcome Trust Centre for Human Genetics, Oxford University, Oxford OX3 7BN, UK. [2]Department of Statistics, Oxford University, Oxford OX1 3TG, UK. [3]The Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. [4]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [5]Department of Genetics, Harvard Medical School, Cambridge, Massachusetts 02142, USA. [6]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SA, UK. [7]Center for Statistical Genetics, Biostatistics, University of Michigan, Ann Arbor, Michigan 48109, USA. [8]Illumina United Kingdom, Chesterford Research Park, Little Chesterford, Near Saffron Walden, Essex CB10 1XL, UK. [9]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. [10]Center for Comparative and Population Genomics, Cornell University, Ithaca, New York 14850, USA. [11]Department of Genome Sciences, University of Washington School of Medicine and Howard Hughes Medical Institute, Seattle, Washington 98195, USA. [12]European Bioinformatics Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SD, UK. [13]Baylor College of Medicine, Human Genome Sequencing Center, Houston, Texas 77030, USA. [14]US National Institutes of Health, National Human Genome Research Institute, 31 Center Drive, Bethesda, Maryland 20892, USA. [15]Centre of Genomics and Policy, McGill University, Montréal, Québec H3A 1A4, Canada. [16]European Molecular Biology Laboratory, Genome Biology Research Unit, Meyerhofstraße 1, 69117 Heidelberg, Germany. [17]Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. [18]Max Planck Institute for Molecular Genetics, Ihnestraße 63-73, 14195 Berlin, Germany. [19]Dahlem Centre for Genome Research and Medical Systems Biology, D-14195 Berlin-Dahlem, Germany. [20]The Genome Center, Washington University School of Medicine, St Louis, Missouri 63108, USA. [21]Department of Biology, Boston College, Chestnut Hill, Massachusetts 02467, USA. [22]Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington 98195, USA. [23]Affymetrix, Inc., Santa Clara, California 95051, USA. [24]US National Institutes of Health, National Center for Biotechnology Information, 45 Center Drive, Bethesda, Maryland 20892, USA. [25]BGI-Shenzhen, Shenzhen 518083, China. [26]The Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, DK-2200 Copenhagen, Denmark. [27]Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark. [28]Alacris Theranostics GmbH, D-14195 Berlin-Dahlem, Germany. [29]Department of Genetics, Albert Einstein College of Medicine, Bronx, New York 10461, USA. [30]Department of Computational Medicine and Bioinfomatics, University of Michigan, Ann Arbor, Michigan 48109, USA. [31]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA. [32]Seaver Autism Center and Department of Psychiatry, Mount Sinai School of Medicine, New York, New York 10029, USA. [33]Department of Nanobiomedical Science, Dankook University, Cheonan 330-714, South Korea. [34]Department of Biological Sciences, Dankook University, Cheonan 330-714, South Korea. [35]Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, New York 14853, USA. [36]Center for Systems Biology and Department Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA. [37]Institute of Medical Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK. [38]Illumina, Inc., San Diego, California 92122, USA. [39]Molecular Epidemiology Section, Department of Medical Statistics and Bioinformatics, Leiden University Medical Center 2333 ZA, The Netherlands. [40]Department of Biological Sciences, Louisiana State University, Baton

Rouge, Louisiana 70803, USA. [41]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [42]Department of Anthropology, Penn State University, University Park, Pennsylvania 16802, USA. [43]Department of Genetics, Stanford University, Stanford, California 94305, USA. [44]Ancestry.com, San Francisco, California 94107, USA. [45]Blavatnik School of Computer Science, Tel-Aviv University, 69978 Tel Aviv, Israel. [46]Department of Microbiology, Tel-Aviv University, 69978 Tel Aviv, Israel. [47]International Computer Science Institute, Berkeley, California 94704, USA. [48]The Translational Genomics Research Institute, Phoenix, Arizona 85004, USA. [49]Life Technologies, Beverly, Massachusetts 01915, USA. [50]Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, California 90024, USA. [51]Department of Psychiatry, University of California, San Diego, La Jolla, California 92093, USA. [52]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, California 92093, USA. [53]Department of Computer Science, University of California, San Diego, La Jolla, California 92093, USA. [54]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, New York 10461, USA. [55]Department of Bioengineering and Therapeutic Sciences and Medicine, University of California, San Francisco, California 94158, USA. [56]Center for Biomolecular Science and Engineering, University of California, Santa Cruz, California 95064, USA. [57]Howard Hughes Medical Institute, Santa Cruz, California 95064, USA. [58]Department of Human Genetics, University of Chicago, Chicago, Illinois 60637, USA. [59]Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK. [60]Department of Genetic Medicine and Development, University of Geneva Medical School, 1211 Geneva, Switzerland. [61]Institute for Genetics and Genomics in Geneva (iGE3), University of Geneva, 1211 Geneva, Switzerland. [62]Swiss Institute of Bioinformatics, 1211 Geneva, Switzerland. [63]Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA. [64]IST/High Performance and Research Computing, University of Medicine and Dentistry of New Jersey, Newark, New Jersey 07107, USA. [65]Department of Invertebrate Zoology, American Museum of Natural History, New York, New York 10024, USA. [66]Istituto di Ricerca Genetica e Biomedica, CNR, Monserrato, 09042 Cagliari, Italy. [67]Department of Anthropology, University of Michigan, Ann Arbor, Michigan 48109, USA. [68]Dipartimento di Scienze Biomediche, Università delgi Studi di Sassari, 07100 Sassari, Italy. [69]Center for Advanced Studies, Research, and Development in Sardinia (CRS4), AGCT Program, Parco Scientifico e tecnologico della Sardegna, 09010 Pula, Italy. [70]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [71]University of Michigan Sequencing Core, University of Michigan, Ann Arbor, Michigan 48109, USA. [72]National Institute on Aging, Laboratory of Genetics, Baltimore, Maryland 21224, USA. [73]Department of Pediatrics, University of Montréal, Ste. Justine Hospital Research Centre, Montréal, Québec H3T 1C5, Canada. [74]Department of Biology, University of Puerto Rico, Mayagüez, Puerto Rico 00680, USA. [75]The University of Texas Health Science Center at Houston, Houston, Texas 77030, USA. [76]Eccles Institute of Human Genetics, University of Utah School of Medicine, Salt Lake City, Utah 84112, USA. [77]Department of Genetics, Rutgers University,The State University of New Jersey, Piscataway, New Jersey 08854, USA. [78]Department of Medicine, Division of Medical Genetics, University of Washington, Seattle, Washington 98195, USA. [79]Department of Computer Engineering, Bilkent University, TR-06800 Bilkent, Ankara, Turkey. [80]Department of Computer Science, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada. [81]Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77230, USA. [82]Department of Haematology, University of Cambridge and National Health Service Blood and Transplant, Cambridge CB2 1TN, UK. [83]Institute of Genetics and Biophysics, National Research Council (CNR), 80125 Naples, Italy. [84]Program in Computational Biology and Bioinformatics, Yale University, New Haven, Connecticut 06520, USA. [85]Department of Computer Science, Yale University, New Haven, Connecticut 06520, USA. [86]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, USA. [87]Department of Chemistry, Yale University, New Haven, Connecticut 06520, USA. [88]Beyster Center for Genomics of Psychiatric Diseases, University of California, San Diego, La Jolla, California 92093, USA. [89]US National Institutes of Health, National Human Genome Research Institute, 50 South Drive, Bethesda, Maryland 20892, USA. [90]Division of Allergy and Clinical Immunology, School of Medicine, Johns Hopkins University, Baltimore, Maryland 21205, USA. [91]Coriell Institute for Medical Research, Camden, New Jersey 08103, USA. [92]Centre for Health, Law and Emerging Technologies, University of Oxford, Oxford OX3 7LF, UK. [93]Genetic Alliance, London N1 3QP, UK. [94]Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. [95]Department of Medical History and Bioethics, Morgridge Institute for Research, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA. [96]University of Wisconsin Law School, Madison, Wisconsin 53706, USA. [97]The Ethox Centre, Department of Public Health, University of Oxford, Old Road Campus, Oxford OX3 7LF, UK. [98]US National Institutes of Health, Center for Research on Genomics and Global Health, National Human Genome Research Institute, 12 South Drive, Bethesda, Maryland 20892, USA. [99]Institute for Genome Sciences and Policy, Duke University, Durham, North Carolina 27708, USA. [100]Department of Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA. [101]Department of Animal Biology, Unit of Anthropology, University of Barcelona, 08028 Barcelona, Spain. [102]Cancer and Immunogenetics Laboratory, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, UK. [103]Laboratory of Molecular Genetics, Institute of Biology, University of Antioquia, Medellín, Colombia. [104]Peking University Shenzhen Hospital, Shenzhen 518036, China. [105]Institute of Medical Biology, Chinese Academy of Medical Sciences and Peking Union Medical College, Kunming 650118, China. [106]Instituto de Biologia Molecular y Celular del Cancer, Centro de Investigacion del Cancer/IBMCC (CSIC-USAL), Institute of Biomedical Research of Salamanca (IBSAL) & Banco Nacional de ADN Carlos III, University of Salamanca, 37007 Salamanca, Spain. [107]Instituto de Biologia Molecular y Celular del Cancer, Centro de Investigacion del Cancer/IBMCC (CSIC-USAL), Institute of Biomedical Research of Salamanca (IBSAL) & Cytometry Service and Department of Medicine, 37007 University of Salamanca, Salamanca, Spain. [108]Ponce School of Medicine and Health Sciences, Ponce, Puerto Rico 00716, USA. [109]US National Institutes of Health, National Human Genome Research Institute, 5635 Fishers Lane, Bethesda, Maryland 20892, USA. [110]Wellcome Trust, Gibbs Building, 215 Euston Road, London NW1 2BE, UK.

‡Deceased.

doi:10.1038/nature11525

Generation of functional thyroid from embryonic stem cells

Francesco Antonica[1], Dominika Figini Kasprzyk[1], Robert Opitz[1], Michelina Iacovino[2], Xiao-Hui Liao[3], Alexandra Mihaela Dumitrescu[3], Samuel Refetoff[3,4], Kathelijne Peremans[5], Mario Manto[6], Michael Kyba[2] & Sabine Costagliola[1]

The primary function of the thyroid gland is to metabolize iodide by synthesizing thyroid hormones, which are critical regulators of growth, development and metabolism in almost all tissues. So far, research on thyroid morphogenesis has been missing an efficient stem-cell model system that allows for the *in vitro* recapitulation of the molecular and morphogenic events regulating thyroid follicular-cell differentiation and subsequent assembly into functional thyroid follicles. Here we report that a transient overexpression of the transcription factors NKX2-1 and PAX8 is sufficient to direct mouse embryonic stem-cell differentiation into thyroid follicular cells that organize into three-dimensional follicular structures when treated with thyrotropin. These *in vitro*-derived follicles showed appreciable iodide organification activity. Importantly, when grafted *in vivo* into athyroid mice, these follicles rescued thyroid hormone plasma levels and promoted subsequent symptomatic recovery. Thus, mouse embryonic stem cells can be induced to differentiate into thyroid follicular cells *in vitro* and generate functional thyroid tissue.

The mammalian thyroid consists of two endocrine cell types, the thyroid follicular cells (TFCs) that produce the thyroid hormones T3 and T4 and the C-cells that produce calcitonin[1]. In the adult thyroid gland, TFCs are organized into follicular structures[2], in which a monolayer of polarized TFCs enclose a luminal compartment filled with a colloidal mass containing thyroid hormone precursors bound to thyroglobulin[3]. A follicular organization of TFCs is considered to be the prerequisite for efficient thyroid hormone biosynthesis[4]. It has been demonstrated that NKX2-1 (ref. 5) and PAX8 (ref. 6) function are vital for TFC survival, differentiation[2] and function during thyroid organogenesis and in mature thyroid tissue[1]. During thyroid organogenesis, the onset of NKX2-1 (ref. 7) and PAX8 (ref. 8) co-expression in a small group of ventral foregut endodermal cells represents the first molecular marker of cell specification towards a TFC fate. Although NKX2-1 (ref. 7) and PAX8 (ref. 8) are expressed individually in a variety of tissues and cell types, their co-expression is restricted to cells committed to differentiate into TFCs. Induced overexpression of defined transcription factors has been shown to have a directing effect on the differentiation of embryonic stem cells (ESCs) into specific cell types[9–11]. Despite the success of this experimental approach for cell differentiation or reprogramming, protocols promoting coordinated self-assembly of differentiated cells into distinct morphological units with functional properties reminiscent of organs and tissues *in vivo*[12–14] are still very sparse. In this study, we explore whether overexpression of the transcription factors NKX2-1 and PAX8 could promote differentiation of murine ESCs into TFCs and subsequent self-formation of thyroid follicles.

In vitro thyroid cell differentiation

Because the factors and signalling pathways inducing concurrent expression of NKX2-1 and PAX8 have not yet been resolved[2], we generated recombinant ESC lines (Fig. 1a and Supplementary Fig. 1a, c, e) in which expression of these transcription factors can be temporally induced on the addition of doxycycline (Dox; 1 $\mu g\,ml^{-1}$) to the medium[15] (Supplementary Fig. 1b, d, f). These genetic manipulations did not affect the pluripotent state of the ESCs (Supplementary Fig. 1g). In our experimental set-up, Dox induction of NKX2-1 and PAX8 was initiated after a 4-day ESC culture in hanging drops to allow for differentiation into embryoid bodies (Fig. 1b). We first used a recombinant ESC line in which Dox treatment induces NKX2-1 and PAX8 overexpression (Supplementary Fig. 1a, b). After 3 days of Dox treatment (on days 4, 5 and 6), co-expression of NKX2-1 and PAX8 was detectable by immunofluorescence in almost all Dox-treated cells on day 7 but never in cells incubated in the absence of Dox (Supplementary Fig. 2a). To determine whether the combined activity of NKX2-1 and PAX8 promotes TFC differentiation, we first examined the expression of various TFC markers by quantitative reverse transcriptase PCR (qRT–PCR). Notably, messenger RNA expression of functional markers, including the thyroid-stimulating hormone (TSH) receptor (*Tshr*), the sodium/iodide symporter NIS (*Slc5a5*) and thyroglobulin (*Tg*), was strongly upregulated within 3 days of Dox treatment (Supplementary Fig. 2b). The expression of *Foxe1*, another key transcription factor for thyroid development[16], was also upregulated (Supplementary Fig. 2b), and NKX2-1$^+$ FOXE1$^+$ cells were prominent throughout cell cultures (Supplementary Fig. 2c). Interestingly, our qRT–PCR analyses also demonstrated a robust increase in endogenous *Nkx2-1* and *Pax8* mRNA levels (Supplementary Fig. 2d), indicating an auto-induction of these transcription factors. Together, these data demonstrate that forced co-expression of NKX2-1 and PAX8 readily acts on cell fate, driving the differentiation towards a TFC lineage. However, assembly of Dox-treated cells into three-dimensional aggregates reminiscent of follicle-like epithelial structures was rarely observed under these conditions. This was true for cell cultures using a variety of distinct Dox treatment protocols, suggesting that additional factors might be required to promote follicular morphogenesis. We therefore revised the treatment protocol on the basis of two critical observations. First, we limited the Dox treatment to a 3-day period from day 4 to day 6 (Fig. 1b), as this seemed to be sufficient to induce the auto-induction of

---

[1]Institute of Interdisciplinary Research in Molecular Human Biology (IRIBHM), Université Libre de Bruxelles, 808 route de Lennik, 1070 Brussels, Belgium. [2]Lillehei Heart Institute and Department of Pediatrics, University of Minnesota, Minneapolis, 55455 Minnesota, USA. [3]Department of Medicine, The University of Chicago, Chicago, 606370 Illinois, USA. [4]Departments of Pediatrics & Genetics, The University of Chicago, Chicago, 60637 Illinois, USA. [5]Department of Veterinary Medical Imaging & Small Animal Orthopaedics, Faculty of Veterinary Medicine, Ghent University, B-9820 Merelbeke, Belgium. [6]FNRS, ERASME, Université Libre de Bruxelles, 808 Route de Lennik, 1070 Brussels, Belgium.

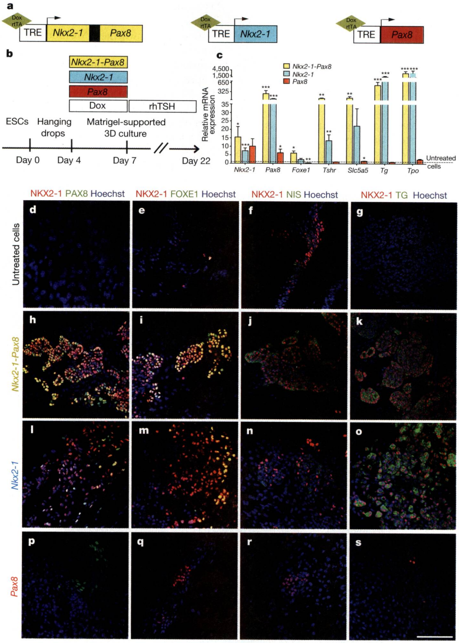


Figure 1 | Ectopic expression of *Nkx2-1* and *Pax8* promotes the differentiation of ESCs into thyroid follicles. **a**, Schematic representation of tetracycline-inducible murine ESC lines. **b**, Schematic diagram of the thyroid follicle differentiation protocol from ESCs. **c**, Expression of endogenous *Nkx2-1* and *Pax8*, *Foxe1*, *Tshr*, *Slc5a5*, *Tg* and *Tpo* at day 22 in cells differentiated after Dox-mediated induction of *Nkx2-1-Pax8* (yellow columns), *Nkx2-1* (cyan columns) and *Pax8* (red columns). Relative expression of each transcript is presented as fold change compared to untreated cells at day 22 as mean ± s.e.m. ($n = 6$). Unpaired t-test was used for statistical analysis. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. **d–s**, Immunostaining at day 22 of untreated cells (**d–g**) and after Dox-mediated induction of *Nkx2-1-Pax8* (**h–k**), *Nkx2-1* (**l–o**) and *Pax8* (**p–s**) for NKX2-1 and PAX8 (**d, h, l, p**), NKX2-1 and FOXE1 (**e, i, m, q**), NKX2-1 and NIS (**f, j, n, r**) and NKX2-1 and TG (**g, k, o, s**). Scale bars, 100 µm. rtTA, reverse tetracycline transactivator; TRE, tetracycline-responsive element.

endogenous *Nkx2-1* and *Pax8* mRNA expression. Second, we treated cells from day 7 onwards with recombinant human TSH (rhTSH; 1 mU ml^{-1}) (Fig. 1b), as the robust upregulation of *Tshr* mRNA indicated that the cells had acquired an ability to respond to rhTSH. When using this sequential Dox–rhTSH treatment schedule, qRT-PCR analyses of day-22 cell cultures indeed showed a sustained elevation of endogenous *Nkx2-1* and *Pax8* mRNA expression (Fig. 1c and Supplementary Fig. 3a, b). Moreover, these cell cultures also showed a high expression of functional TFC markers such as *Tshr*, *Slc5a5*, *Tg* and *Tpo* (which codes for thyroid peroxidase) (Fig. 1c and Supplementary Fig. 3a, b). Immunofluorescence analyses at day 22 demonstrated that the sequential Dox–rhTSH treatment clearly promoted the differentiation of NKX2-1^{+} cells co-expressing PAX8 (Fig. 1h and Supplementary Fig. 3g, k), FOXE1 (Fig. 1i and Supplementary Fig. 3h, l), NIS (Fig. 1j; Supplementary Fig. 3i, m) and TG (Fig. 1k and Supplementary Fig. 3j, n). No such co-expression was seen in the absence of Dox-induced transgene expression (Fig. 1d–g and Supplementary Fig. 3c–f). The efficiency of TFC differentiation, quantified as the percentage of NKX2-1^{+} PAX8^{+} cells, reached 60.5 ± 8.1% in the sequential Dox–rhTSH treatment protocol (mean ± s.e.m., $n = 3$; see Methods). Most importantly, and in contrast to a 3-day Dox treatment without subsequent TSH treatment, the sequential Dox–rhTSH treatment greatly stimulated the assembly of TFC-like cells into distinct three-dimensional structures reminiscent of thyroid follicles (Supplementary Fig. 3k–n; compare with Supplementary Fig. 3g–j).

In vitro formation of functional follicles

A follicular organization of TFCs is considered the prerequisite for thyroid hormone biosynthesis, which occurs under physiological conditions extracellularly at the TFC–colloid interface. Iodide organification requires a complex biosynthetic machinery including NIS-mediated iodide uptake[17] at the basal pole[18], TG synthesis and targeting to the apical pole, H_2O_2 generation at the TFC–colloid interface by dual oxidase and TPO-mediated iodination of TG[4] (Fig. 2a). Immunofluorescence analyses of the follicular aggregates demonstrated polarization characteristics consistent with thyroid follicles in intact animals (Fig. 2b). The NKX2-1$^+$ cells surrounding a luminal space showed basolateral localization of NIS (Fig. 2c) and E-cadherin (Fig. 2d), as well as apical localization of zona occludens 1 (ZO-1; also known as TJP1) (Fig. 2e). Conversely, TG immunofluorescence was observed both intracellularly and in the luminal compartment (Fig. 2f). It should be noted that neither endothelial cells nor C-cells were observed after 22 days of cell culture, as judged by the absence of detectable PECAM-1 and calcitonin staining, respectively. Together, these data demonstrate the efficiency of the sequential Dox-rhTSH treatment to generate cells with a molecular signature highly similar to TFCs and to promote the assembly of these TFC-like cells into three-dimensional follicular structures strongly resembling thyroid follicles. We next studied whether Dox induction of either NKX2-1 or PAX8 alone was sufficient to promote TFC differentiation and follicle morphogenesis. Dox induction of NKX2-1 alone for 3 days was sufficient to upregulate the expression of various TFC markers as assessed by qRT–PCR (Supplementary Fig. 4a, b) and by immunostaining for PAX8 (Supplementary Fig. 4c) and FOXE1 (Supplementary Fig. 4d). However, with the exception of *Pax8* mRNA expression, the effects of NKX2-1 overexpression were weaker relative to Dox induction of NKX2-1 and PAX8 (Supplementary Fig. 2b) and no upregulation was evident for endogenous *Nkx2-1* mRNA (Supplementary Fig. 4a). When using the sequential Dox–rhTSH protocol, NKX2-1 induction was clearly sufficient to promote differentiation towards a TFC-like cell fate (Fig. 1c, l–o) but failed to efficiently promote formation of follicle-like cell aggregates (compare Fig. 1k with Fig. 1o). Given the vital role of TSH treatment for folliculogenesis, as observed in the NKX2-1- and PAX8-overexpression model, the comparatively low level of *Tshr* upregulation induced by overexpression of NKX2-1 might represent one plausible factor explaining the lack of follicle morphogenesis in the

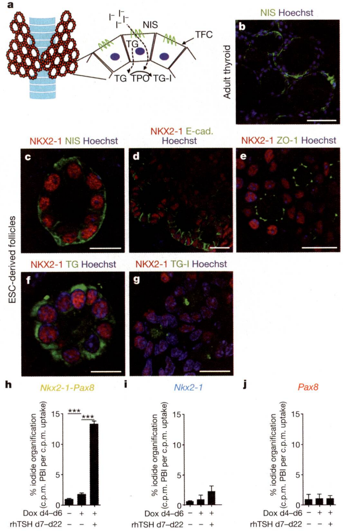

a
I⁻ I⁻ I⁻
NIS
TFC
TG
TG TPO TG-I
NIS Hoechst
b
Adult thyroid
NKX2-1 NIS Hoechst
c
NKX2-1 E-cad. Hoechst
d
NKX2-1 ZO-1 Hoechst
e
ESC-derived follicles
NKX2-1 TG Hoechst
f
NKX2-1 TG-I Hoechst
g
h
Nkx2-1-Pax8
% iodide organification (c.p.m. PBI per c.p.m. uptake)
15
10
5
0
***
***
Dox d4–d6 − + +
rhTSH d7–d22 − − +
i
Nkx2-1
% iodide organification (c.p.m. PBI per c.p.m. uptake)
15
10
5
0
Dox d4–d6 − + +
rhTSH d7–d22 − − +
j
Pax8
% iodide organification (c.p.m. PBI per c.p.m. uptake)
15
10
5
0
Dox d4–d6 − + +
rhTSH d7–d22 − − +

Figure 2 | ESC-derived thyroid cells show full morphological and functional maturation. a, Schematic diagram of the thyroid gland organized in follicles. **b**, Immunostaining of NIS in adult thyroid tissue. **c–f**, Immunofluorescence at day 22 of thyroid follicles derived from ESCs on ectopic expression of *Nkx2-1* and *Pax8* for NKX2-1 and NIS (**c**), NKX2-1 and E-cadherin (E-cad.) (**d**), NKX2-1 and ZO-1 (**e**) and NKX2-1 and TG (**f**). **g**, Immunodetection of TG-I in the luminal compartment of NKX2-1-positive follicles. **h–j**, Iodide-organification assay in cells differentiated after Dox induction of *Nkx2-1-Pax8* (**h**), *Nkx2-1* (**i**) and *Pax8* (**j**). Histograms show the organification percentage of iodine-125 at day 22 in cells differentiated without Dox and rhTSH (left column), in the presence of Dox only (centre column) and on Dox and rhTSH treatment (right column). Data are mean ± s.e.m. ($n = 3$). Tukey's multiple comparison test was used for statistical analysis. $***P < 0.001$. Scale bars, 200 μm (**b**) and 20 μm (**c–g**). PBI, protein-bound ^{125}I.

NKX2-1-overexpression model (see Fig. 1c and Supplementary Fig. 4a, b). In notable contrast to NKX2-1 overexpression, Dox induction of PAX8 had only marginal effects on TFC marker expression, both at the mRNA (Supplementary Fig. 5a, b) and protein-expression level (Supplementary Fig. 5c, d). Treatment with rhTSH after Dox induction of PAX8 did not promote TFC differentiation as judged from the day-22 analyses of TCF marker gene expression (Fig. 1c) and immunofluorescence staining (Fig. 1p–s). Together, these data indicate that transient NKX2-1 overexpression is sufficient to drive cell differentiation towards the TFC lineage. Conversely, PAX8 did not promote TFC differentiation when overexpressed alone but strongly enhanced TFC differentiation when overexpressed together with NKX2-1. Most importantly, overexpression of both NKX2-1 and PAX8 was a critical requirement for the rhTSH-dependent self-assembly of TFC into follicle-like structures. The high similarity between *in vitro*-generated follicle-like structures and true thyroid follicles, both at the molecular and morphological level, prompted us to examine whether these follicular structures were functional for the two hallmarks of thyroid tissue: iodide trapping and thyroid hormone synthesis. We therefore assessed day-22 cell cultures derived from the sequential Dox–rhTSH protocol for their capacity to organify iodide. The first evidence for active iodide organification was obtained by immunofluorescence detection of iodinated TG (TG-I) within the luminal compartment of follicular aggregates (Fig. 2g) using an antibody that selectively recognizes iodinated TG epitopes[19]. Positive TG-I staining was limited to cell cultures obtained after Dox induction of NKX2-1 and PAX8 and subsequent rhTSH treatment. We next used a 'classical' iodide-organification assay that measured the relative incorporation of radioiodine into TCA-precipitable proteins after 2 h of incubation in a ^{125}I-supplemented medium. Measurements of radioiodine incorporation corroborated the TG-I staining results as a strong and significant increase of iodide organification was exclusively seen in cell cultures on Dox induction of NKX2-1 and PAX8 and subsequent rhTSH treatment (Fig. 2h–j and Supplementary Fig. 6). The capacity of such cell cultures for robust iodide organification is in line with their TFC-like molecular signature and their three-dimensional follicular organization. In turn, the lack of similar functional properties of cell cultures derived after the induction of either NKX2-1 (Fig. 2i) or PAX8 alone (Fig. 2j) would be consistent with a failure of TFC differentiation (Dox induction of PAX8 alone) or a reduced competence in forming follicular aggregates on rhTSH treatment (Dox induction of NKX2-1 alone). Together, our data demonstrate that the differentiation protocol relying on overexpression of NKX2-1 and PAX8 allows for deriving *in vitro* follicular organoid structures that recapitulate molecular, morphological and functional properties of bona fide thyroid follicles. In addition, this protocol highlights the vital role of TSH in completing the process of follicular maturation.

In vivo functionality of derived follicles

To assess the potential *in vivo* functionality of the ESC-derived thyroid follicles, we grafted follicular organoids under the kidney capsules of female mice previously made hypothyroid by intraperitoneal ^{131}I injection (Fig. 3a). Histological evaluation of the kidney region 1 month after transplantation demonstrated successful integration of grafted organoids in the host niche (Fig. 3b, c). At the grafting site (Fig. 3b), numerous follicles containing a monolayered epithelium were present at the cortical area of the host organ (Fig. 3c and Supplementary Fig. 7a). *In situ* chromosome Y painting on paraffin sections of the grafting region provided clear evidence that the follicular structures originated from male ESC-derived organoids, on a background of female kidney tissue of the hosts (Supplementary Fig. 8a, b). Immunostaining demonstrated that the follicular epithelium was made up of cells positive for NKX2-1 (Fig. 3d), PAX8 (Fig. 3e) and FOXE1 (Fig. 3f), a molecular signature highly similar to the thyrocytes of orthotopic thyroid tissue. Further immunohistochemical analyses corroborated the development of functional thyroid follicles at the grafting sites, including cytosolic TG expression and TG deposition in the luminal compartment (Fig. 3g), polarized NIS (Fig. 3h) and E-cadherin (Supplementary Fig. 7b) expression at the basolateral membrane as well as detection of the thyroid hormone T4 in the colloid (Fig. 3i and Supplementary Fig. 7a). Importantly, immunostaining for the pan-endothelial marker PECAM-1 showed that thyroid follicles were surrounded by a dense network of small blood vessels, demonstrating the formation of classical angio-follicular units (Supplementary Fig. 7c, d). Lastly, we could not detect any calcitonin staining in the grafted tissue, whereas calcitonin staining was clearly detectable in the orthotopic thyroid tissue of adult mice (data not shown). Thus, consistent with the proposed origin and migration path of C-cells[20], the ectopic thyroid tissue grafts are free of C-cells, indicating that our differentiation protocol does not promote C-cell development.

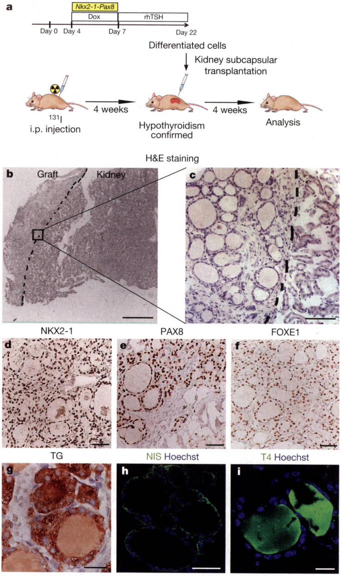


Figure 3 | Grafting of ESC-derived thyroid follicles in mice. **a**, Schematic diagram of protocol for ESC-derived thyroid follicle transplantation in the renal capsule of mice with radio-iodine-ablated thyroid (hypothyroid mice). **b–i**, Histological analysis of kidney sections 4 weeks after grafting. Hematoxylin and eosin staining on optimal cutting temperature embedded grafted kidney showed the localization of the transplanted tissue in the cortical area of the host organ (left side) (**b**) and single cuboidal epithelium organization of transplanted tissue (**c**), the immunohistochemistry of NKX2-1 (**d**), PAX8 (**e**), FOXE1 (**f**), TG (**g**), and the immunofluorescence of NIS (**h**) and T4 (**i**) in the grafted tissue. i.p., intraperitoneal. Scale bars, 300 μm (**b**), 100 μm (**c**), 50 μm (**d**, **e**, **f**, **h**) and 20 μm (**g**, **i**).

Functional rescue in hypothyroid mice

We next evaluated the ability of the ESC-derived tissue grafts to restore thyroid homeostasis in mice with radioiodine-ablated thyroid tissue. Plasma T4 levels measured in female mice at 1 month after ^{131}I injection showed a severe hypothyroid status at the time when ESC-derived organoids were transplanted (Fig. 4a and Supplementary Fig. 9a). Four weeks after grafting (8 weeks after ^{131}I injection), mice grafted with cells that were differentiated on overexpression of NKX2-1 and PAX8 and subsequent rhTSH treatment showed a substantial increase in plasma T4 levels (Fig. 4b and Supplementary Fig. 9b), with a complete rescue of thyroid homeostasis being evident in eight out of nine animals. Notably, mice transplanted with cells that were differentiated without Dox and rhTSH treatment remained hypothyroid (Fig. 4b), as was the case for mice that received no grafts at all (Supplementary Fig. 9b). To demonstrate that the grafted thyroid tissue was responsible for the restoration of plasma T4 levels, we performed whole-body scintigraphy of mice after intramuscular injection of 99mTc-pertechnetate, a γ emitter transported by the sodium/iodide symporter (Fig. 4c, d). As shown in Fig. 4c, strong 99mTc-pertechnetate uptake was observed in control mice in the neck region (where the thyroid and salivary glands reside), the stomach and the bladder. In athyroid mice grafted with ESC-derived thyroid follicles, 99mTc-pertechnetate uptake was markedly decreased in the neck region, owing to the absence of the thyroid gland (Fig. 4d). The remaining weak signal in the neck region was due to NIS activity in the salivary gland. Importantly, a very strong signal was detectable at the grafting site close to the kidney. These data provide strong evidence that the grafted tissue was responsible for the restoration of plasma T4 levels. Along with the restoration of normal T4 plasma levels, mice grafted with ESC-derived thyroid follicles also show a progressive decrease in plasma TSH levels (Fig. 4e and Supplementary Fig. 9c). Moreover, acute TSH administration[21] to athyroid grafted mice was effective in producing an increase in circulating levels of T4, suggesting TSH responsiveness of the grafted tissue (Supplementary Fig. 10). To examine whether our grafting approach also resulted in a symptomatic recovery, we analysed body temperature in mice from the different treatment groups. Decreased body temperature was found to be a robust and sensitive response to lowered plasma thyroid hormone concentrations. Importantly, mice grafted with ESC-derived thyroid follicles showed a full normalization of body temperature at 4 weeks after transplantation, providing a compelling example for symptomatic recovery along with the normalization of plasma hormone concentrations (Fig. 4f and Supplementary Fig. 11). These *in vivo* data clearly demonstrate that ESC-derived thyroid follicles have potent functional capacity to compensate for the lack of orthotopic thyroid tissue, allowing for the full rescue of experimentally induced hypothyroidism (Fig. 4g).

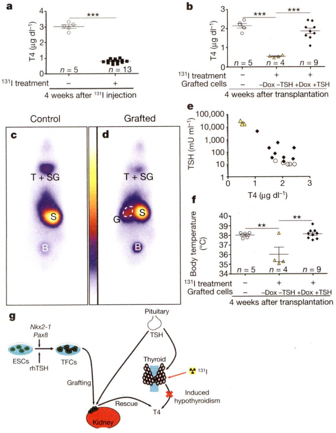


Figure 4 | Rescue of experimentally induced hypothyroidism by transplantation of ESC-derived thyroid follicles. **a**, Total plasma T4 levels 4 weeks after injection in untreated mice (open circles) and iodine-131-treated mice (black squares). **b**, Total plasma T4 levels 4 weeks after the transplantation of differentiated cells in iodine-131-treated mice. **c**, **d**, Whole-body images of mice 30 min after the injection of 99mTc-pertechnetate. Four weeks after grafting, a body scan was performed on untreated control mice (**c**) or iodine-131-treated mice grafted with ESC-derived follicles (**d**). B, bladder; G, grafted ESC-derived follicles; S, stomach; T + SG, thyroid and salivary glands. **e**, Relationship between plasma TSH and T4 levels 4 weeks after grafting. **f**, Body-temperature measurements 4 weeks after grafting. In **b**, **e**, **f**, open circles show iodine-131-untreated and ungrafted mice; yellow triangles show mice treated with iodine-131 and grafted with cells differentiated without Dox and rhTSH (−Dox −TSH) and black diamonds show mice treated with iodine-131 and grafted with cells differentiated with Dox and rhTSH (+Dox +TSH). The values are shown as a dot plot (**a**, **b**, **f**) or scatter plot (**e**) and data are mean ± s.e.m. Unpaired *t*-test (**a**) and Tukey's Multiple Comparison Test (**b**, **f**) were used for statistical analysis. $**P < 0.01$, $***P < 0.001$. **g**, Summary diagram showing that *Nkx2-1* and *Pax8* co-expression in combination with rhTSH treatment leads to the differentiation of ESCs into fully functional thyroid follicles that promote *in vivo* hormonal and symptomatic recovery of the hypothyroid state.

Conclusion and perspectives

We have developed a protocol that allows for the generation of functional thyroid follicles from ESCs on the basis of transient overexpression of two transcription factors followed by rhTSH treatment. Recently, elegant studies have demonstrated self-formation of two ectoderm-derived complex organs, adenohypophysis[13] and optic cup[12], using three-dimensional ESC culture systems, as well as the capacity of colon stem cells to recapitulate *in vitro* the self-organization of an endoderm-derived tissue, the crypt–villus structures[22,23]. Although a few previous studies have reported on the detection of thyrocyte-like cells in ESC cultures[24–27], the present study is the first to demonstrate self-formation of thyroid follicles from ESC-derived TFCs and their capacity for iodide organification *in vitro*. Importantly, when transplanted into mice, the ESC-derived cells generated functional thyroid tissue able to rescue thyroid hormone deficits in athyroid animals. The latter finding of our study in particular opens a new avenue for application of stem-cell technologies in the treatment of hypothyroidism, an area that has so far received relatively little attention in regenerative medicine. In this context, one should bear in mind that congenital hypothyroidism, resulting from either dysfunctional (15%) or dysplastic (85%) thyroid tissue, is the most common congenital endocrine disease in humans, affecting one in 2,000 newborns[28].

METHODS SUMMARY

Recombinant murine ESC (A2Lox *Nkx2-1-Pax8*, A2Lox *Nkx2-1* and A2Lox *Pax8*) lines, generated as previously described in ref. 15, were differentiated in embryoid bodies by the hanging drop method[29]. Embryoid bodies were embedded in growth-factor-reduced Matrigel and re-plated into 12-well plates. The subsequent exposure to Dox and rhTSH was performed as described in Fig. 1b. During *in vitro* differentiation, cells were subjected to extensive phenotypic characterization using qRT–PCR, immunohistochemistry, immunofluorescence and iodide-organification assays. *In vivo* studies were performed in hypothyroidism-induced mouse models generated as previously described in ref. 30. In cell-transplantation studies, 22-day-long cultures were digested with a dispase–collagenase mixture, and the purified cellular population enriched with ESC-derived thyroid follicles were transplanted under the renal capsule. Histological examination of the kidneys, T4 and TSH plasma levels and body-temperature measurements were performed 4 weeks after transplantation.

Full Methods and any associated references are available in the online version of the paper.

Received 13 March; accepted 17 August 2012.

Published online 10 October 2012.

1. De Felice, M. Thyroid development and its disorders: genetics and molecular mechanisms. *Endocr. Rev.* **25,** 722–746 (2004).
2. De Felice, M. & Di Lauro, R. Minireview: intrinsic and extrinsic factors in thyroid gland development: an update. *Endocrinology* **152,** 2948–2956 (2011).
3. Mauchamp, J., Mirrione, A., Alquier, C. & Andre, F. Follicle-like structure and polarized monolayer: role of the extracellular matrix on thyroid cell organization in primary culture. *Biol. Cell* **90,** 369–380 (1998).
4. Nunez, J. & Pommier, J. Formation of thyroid hormones. *Vitam. Horm.* **39,** 175–229 (1982).
5. Kimura, S. *et al.* The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. *Genes Dev.* **10,** 60–69 (1996).
6. Mansouri, A., Chowdhury, K. & Gruss, P. Follicular cells of the thyroid gland require *Pax8* gene function. *Nature Genet.* **19,** 87–90 (1998).
7. Lazzaro, D., Price, M., de Felice, M. & Di Lauro, R. The transcription factor TTF-1 is expressed at the onset of thyroid and lung morphogenesis and in restricted regions of the foetal brain. *Development* **113,** 1093–1104 (1991).
8. Plachov, D. *et al. Pax8*, a murine paired box gene expressed in the developing excretory system and thyroid gland. *Development* **110,** 643–651 (1990).
9. Kyba, M., Perlingeiro, R. C. & Daley, G. Q. HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell* **109,** 29–37 (2002).
10. Ahfeldt, T. *et al.* Programming human pluripotent stem cells into white and brown adipocytes. *Nature Cell Biol.* **14,** 209–219 (2012).
11. Kamiya, D. *et al.* Intrinsic transition of embryonic stem-cell differentiation into neural progenitors. *Nature* **470,** 503–509 (2011).
12. Eiraku, M. *et al.* Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* **472,** 51–56 (2011).
13. Suga, H. *et al.* Self-formation of functional adenohypophysis in three-dimensional culture. *Nature* **480,** 57–62 (2011).
14. Eiraku, M. & Sasai, Y. Mouse embryonic stem cell culture for generation of three-dimensional retinal and cortical tissues. *Nature Protocols* **7,** 69–79 (2012).
15. Iacovino, M. *et al.* Inducible cassette exchange: a rapid and efficient system enabling conditional gene expression in embryonic stem and primary cells. *Stem Cells* **29,** 1580–1588 (2011).
16. De Felice, M. *et al.* A mouse model for hereditary thyroid dysgenesis and cleft palate. *Nature Genet.* **19,** 395–398 (1998).
17. Dai, G., Levy, O. & Carrasco, N. Cloning and characterization of the thyroid iodide transporter. *Nature* **379,** 458–460 (1996).
18. Carrasco, N. Iodide transport in the thyroid gland. *Biochim. Biophys. Acta* **1154,** 65–82 (1993).
19. Hartog, M. T. D., Boer, M. D., Veenboer, G. J. M. & Vijlder, J. J. M. D. Generation and characterization of monoclonal antibodies directed against noniodinated and iodinated thyroglobulin, among which are antibodies against hormonogenic sites. *Endocrinology* **127,** 3160–3165 (1990).
20. Fontaine, J. Multistep migration of calcitonin cell precursors during ontogeny of the mouse pharynx. *Gen. Comp. Endocrinol.* **37,** 81–92 (1979).
21. Moeller, L. C. *et al.* Hypothyroidism in thyroid transcription factor 1 haploinsufficiency is caused by reduced expression of the thyroid-stimulating hormone receptor. *Mol. Endocrinol.* **17,** 2295–2302 (2003).
22. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche. *Nature* **459,** 262–265 (2009).
23. Yui, S. *et al.* Functional engraftment of colon epithelium expanded *in vitro* from a single adult Lgr5$^+$ stem cell. *Nature Med.* **18,** 618–623 (2012).
24. Lin, R. Y., Kubo, A., Keller, G. M. & Davies, T. F. Committing embryonic stem cells to differentiate into thyrocyte-like cells *in vitro*. *Endocrinology* **144,** 2644–2649 (2003).
25. Arufe, M. C. *et al.* Directed differentiation of mouse embryonic stem cells into thyroid follicular cells. *Endocrinology* **147,** 3007–3015 (2006).
26. Jiang, N. *et al.* Differentiation of E14 mouse embryonic stem cells into thyrocytes *in vitro*. *Thyroid* **20,** 77–84 (2010).
27. Longmire, T. A. *et al.* Efficient derivation of purified lung and thyroid progenitors from embryonic stem cells. *Cell Stem Cell* **10,** 398–411 (2012).
28. Grüters, A. & Krude, H. Detection and treatment of congenital hypothyroidism. *Nature Rev. Endocrinol.* **8,** 104–113 (2012).
29. Bondue, A. *et al.* Mesp1 acts as a master regulator of multipotent cardiovascular progenitor specification. *Cell Stem Cell* **3,** 69–84 (2008).
30. Abel, E. D. *et al.* Divergent roles for thyroid hormone receptor β isoforms in the endocrine axis and auditory system. *J. Clin. Invest.* **104,** 291–300 (1999).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank G. Vassart, C. Blanpain and P. Vanderhaeghen for discussions and comments and V. Janssens for technical help. X.-H.L., A.M.D. and S.R. are supported in part by grants DK15070 and DK91016 from the National Institutes of Health. This work was supported by the Belgian Fonds de la Recherche Scientifique Medicale (FRSM[2]3_4_557_08 and [3]3_4598_12), Action de Recherche Concertée de la Communauté Française de Belgique (ARC N°AUWB-08/13-ULB10), Fonds d'Encouragement à la Recherche and grants from the Belgian National Fund for Scientific Research (FNRS). F.A. and D.F.K. are FNRS and Fund for Research in the Industry and the Agriculture (FRIA) research fellows, R.O. is an FNRS Postdoctoral Researcher and S.C. is an FNRS Senior Research Associate.

Author Contributions F.A. and S.C. developed the project, designed the experiments and analysed the data. F.A. performed most of the *in vitro* experiments and *in vivo* studies. D.F.K. provided technical help for the *in vitro* differentiation and functional characterization of the cells. M.I and M.K. provided A2Lox-Cre embryonic stem cells. X.-H.L. analysed blood TSH levels. A.M.D. read the manuscript and made experimental suggestions. S.R. provided suggestions and advice on the experimental procedures. K.P. performed the whole-body scan. M.M. performed body-temperature measurements. F.A., R.O. and S.C. wrote the manuscript. All authors read and approved the final manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.C. (scostag@ulb.ac.be).

doi:10.1038/nature11549

Spontaneous network formation among cooperative RNA replicators

Nilesh Vaidya[1], Michael L. Manapat[2], Irene A. Chen[3]†, Ramon Xulvi-Brunet[3], Eric J. Hayden[4] & Niles Lehman[1]

The origins of life on Earth required the establishment of self-replicating chemical systems capable of maintaining and evolving biological information. In an RNA world, single self-replicating RNAs would have faced the extreme challenge of possessing a mutation rate low enough both to sustain their own information and to compete successfully against molecular parasites with limited evolvability. Thus theoretical analyses suggest that networks of interacting molecules were more likely to develop and sustain life-like behaviour. Here we show that mixtures of RNA fragments that self-assemble into self-replicating ribozymes spontaneously form cooperative catalytic cycles and networks. We find that a specific three-membered network has highly cooperative growth dynamics. When such cooperative networks are competed directly against selfish autocatalytic cycles, the former grow faster, indicating an intrinsic ability of RNA populations to evolve greater complexity through cooperation. We can observe the evolvability of networks through *in vitro* selection. Our experiments highlight the advantages of cooperative behaviour even at the molecular stages of nascent life.

The 'RNA world' is a plausible stage in the development of life because RNA simultaneously possesses evolvability and catalytic function[1]. An RNA organism that could evolve in such a fashion is likely to have been one of the Earth's first life forms. A search is underway[2,3] for an RNA autoreplicase that relies on its individual genotype to compete for survival and reproduction by Darwinian-type evolution in a fitness landscape. Yet the transition from a prebiotic chemistry to this stage of life is not understood. Several authors have proposed that the most primitive life thrived less on discrete genotypes and instead on collections of molecular types more subject to systems chemistry than to straightforward selection dynamics[4–9]. In particular, it was suggested that webs of functionally linked, genetically related replicators were required in the earliest phases of life's appearance to prevent informational decay (the so-called error catastrophe)[4,10–12].

An empirical demonstration of RNA replicator networks could illuminate critical features of this early stage of life. Ribozymes are good candidates for this because they can evolve outside of an organismal context, construct other RNAs, exhibit self-sustained reproduction, and explore sequence space in efficient ways[13–15]. However, their ability to form catalytic networks capable of expanding as predicted from theory has not yet been shown, despite the observation that collections of nucleic acids have the potential to manifest complexity[6,16]. Simulations show that molecular networks should arise, evolve and provide a population with resistance against parasitic sequences[8]. These results are robust within structured environments such as cells or on grids, but are less so in a solution phase. Recent experimental work *in vitro* has been very successful at demonstrating simple ecologies[17–19], reciprocity between two species[6,16,20], and sustained exponential growth via cross catalysis[15]. Empirical efforts to date have been limited by an inability to expand past reciprocal interactions between two species to prebiotically relevant systems that have the capacity to increase their complexity by expanding to three, and then more, members[18,21]. Specifically, the use of systems in which the recognition domain in the catalyst and the target domain in the substrate are co-located in each replicator has prevented networks of more than two members from forming. If this molecular feature could be circumvented, larger networks could be realized within RNA populations in the test tube and help demonstrate a potential escape from the error catastrophe problem that tends to plague selfish systems.

The *Azoarcus* ribozyme system

The ~200-nucleotide (nt) *Azoarcus* group I intron ribozyme[22] can be broken into fragments that can covalently self-assemble by catalysing recombination reactions in an autocatalytic fashion[23,24] (Supplementary Fig. 1). By allowing variation in the sequence recognition mechanism by which this assembly occurs, which is provided by the 3-nt internal guide sequence (IGS) at the 5′ end of the ribozyme, many such autonomously self-assembling ribozymes become possible. We sought to determine if these ribozymes could display cooperative behaviour if their IGS sequences target the assembly of other ribozymes, but not themselves.

To create a cooperative network, we fragmented the *Azoarcus* ribozyme into two pieces in three different ways with the intent of observing how they could spontaneously reassemble via intermolecular cooperation (Fig. 1a, b). We manipulated the IGS (canonically GUG) and its target triplet to generate both matched and mismatched partners. We mixed various IGS and target pairs in two-piece constructs to test the ability of mismatched pairs to promote self-assembly (Supplementary Fig. 2). From these data, we chose three mismatched pairs that exhibit relatively little autocatalysis: GUG/CGU, GAG/CAU, and GCG/CUU. These crippled pairs are denoted I_1, I_2 and I_3, respectively, meaning that they are informational subsystems, albeit weakly autocatalytic.

We chose the triplet pairs so that when the three subsystems are mixed together, they should constitute a cyclical cooperative network in which the output of one subsystem can catalyse the replication of the next one in the cycle (Fig. 1b). This occurs because the IGS of one subsystem is matched to the target in the next subsystem, and the physical separation of the IGS and its target allows for cycles of more than two members. When the six RNAs (**W**, ***h*•X•Y•Z**, **W•X**, ***h*•Y•Z**, **W•X•Y** and ***h*•Z**; • indicates covalent bonding) are allowed to fold together and be co-incubated in equimolar ratios, we expect the subsystems first to

[1]Department of Chemistry, Portland State University, PO Box 751, Portland, Oregon 97207, USA. [2]School of Engineering and Applied Sciences and Program for Evolutionary Dynamics, Harvard University, Cambridge, Massachusetts 02138, USA. [3]FAS Center for Systems Biology, Harvard University, Cambridge, Massachusetts 02138, USA. [4]Department of Bioengineering, Stanford University, Stanford, California 94305, USA. †Present address: Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106, USA.

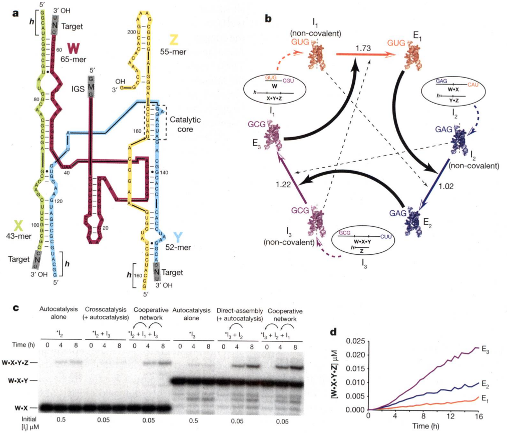


Figure 1 | Cooperative covalent assembly of recombinase ribozymes. **a**, Design of recombinase ribozymes capable of spontaneous cooperative covalent assembly from fragments. The *Azoarcus* ribozyme[25] can be broken at three loop regions to obtain four oligomers capable of self-assembling into a full-length molecule[26,27]. The grey box in **W** (magenta) is the internal guide sequence (IGS), whereas those at the 3′ ends of the **W**, **X** (lime) and **Y** (blue) fragments are recombination targets (tags) recognized by the IGS, which guides the catalysis of a covalent closure (•) of the loops. **b**, A cooperative system comprised of three subsystems, each created from partitioning the molecule into two pieces at different junctions: I_1 (**W** + ***h*•X•Y•Z**), I_2 (**W•X** + ***h*•Y•Z**) and I_3 (**W•X•Y** + ***h*•Z**). Numbers over arrows estimate the cooperative advantage for each step (see text). **c**, Electrophoretic observation of assemblies of E_2 and E_3. The 5′ fragments of I_2 or I_3 were independently 5′-radiolabelled with ^{32}P (that is, $^*I_2$ or $^*I_3$). The reactions were performed by incubating 0.5 μM (for autocatalysis) or 0.05 μM (for direct assembly, cross catalysis and cooperation) of each fragment for 8 h. Where appropriate, the arrows identify the subsystems being assembled by the previous subsystems in the network, where the IGS and recombination tags match. **d**, Yields of individual E_i ribozymes over time, measured every 30 min for 16 h when all six I_i RNA fragments are co-incubated at 0.05 μM.

form non-covalent versions of ribozymes, and then catalyse the formation of covalent versions of the next ribozyme in the cycle.

To test whether cooperation between enzymes occurred, we took several approaches. First, for the cycle to exhibit positive feedback[4], there should be a distinct advantage to being a covalently contiguous ribozyme (E_i), as opposed to remaining fragmented (I_i). Once covalent ribozymes are formed, they should further promote synthesis of their target ribozymes, at faster rates than the non-covalent versions would. When we tested each in isolation, we found that the E_i ribozymes recombined their respective target substrates into products 1.3–6.3-fold more than the I_i versions when assayed separately (Supplementary Fig. 3). Second, by examining each subsystem in isolation or in pairs, we could compare the relative strengths of autocatalysis (E_i synthesizing E_i), cross-catalysis (E_{i+1} synthesizing E_i), and what should be the most efficient, direct catalysis (E_i synthesizing E_{i+1}). When we incubated just the two RNAs from any one subsystem, such as I_2, alone, there is minimal synthesis of the corresponding ribozyme E_2; after a few hours roughly 0.1% of **W•X** is converted into **W•X•Y•Z**. This low background level of autocatalytic synthesis reflects residual catalytic activity available to a mismatched IGS and IGS target, for example GAG with CAU[25], showing that each I_i subsystem has severely limited information-replication potential in isolation. Likewise, when the four RNAs of two subsystems were co-incubated, the cross-catalytic synthesis of the ribozyme corresponding to the preceding subsystem in the cycle is similarly poor, again hindered by an IGS–IGS-target mismatch (Fig. 1c). After only 1 h of incubation, the yield of E_3 from 0.5 μM I_3 is 0.10 ± 0.02% (autocatalysis), and the yield of E_3 from 0.5 μM I_3 and 0.5 μM E_1 is 0.7 ± 0.06% (cross-catalysis), but the yield of E_3 from 0.5 μM I_3 and 0.5 μM I_2 is 13 ± 0.5% (direct catalysis) (data not shown; errors given as s.e.m.). These differences are all statistically significant as measured by *t*-tests several planned comparisons ($P < 0.001$). From these data we determined that direct catalysis is significantly more efficient than catalysis resulting from mismatched IGS sequences and their targets.

When all six RNAs of all three subsystems are co-incubated, cooperation causes the synthesis of **W•X•Y•Z** to rapidly escalate, as expected. The composite yield of full-length RNA after 16 h when I_1, I_2 and I_3 are mixed is 125-fold higher than the sum of the yields of the three subsystems in isolation (Supplementary Fig. 4). This enhancement can be readily visualized after shorter periods of time (Fig. 1c). Each subsystem grows at a different rate (Fig. 1d). The synthesis of E_3 by E_2 is more rapid than that of the other two ribozymes, presumably because the non-covalent version of the enzyme (I_2) is nearly as efficient as the covalent version (E_2); it could also be because certain IGS–IGS target pairs are more efficient[25]. Importantly, we can detect two-step (relayed) cooperativity by comparing the yields with and without the intervening enzyme. In the case of E_1 for example, after 4 h the increase in yield of E_1 upon addition of I_2 to I_1 with I_3 present is 2.5%, whereas the increase in yield of adding I_2 to I_1 without I_3 present is only 0.02%, showing the operation of E_2 through E_3 onto E_1 (Supplementary Table 1); this is supported by doping experiments (Supplementary Fig. 5).

To observe the advantage of cooperation in another way, we constructed a control system in which the I_i molecules could act as catalysts, but could not be covalently assembled themselves because their target sequences were not a match for any catalyst in the system (Supplementary Fig. 5). Cooperation would be manifest when enzymes synthesize other enzymes, and there is some benefit to being covalent. Thus we measured the yields of **W•X•Y•Z** molecules at 8 h in this control system and in our normal system (that is, Fig. 1b). The yields in the control system were consistently worse, and we calculated the ratio (E_i catalysis + I_i catalysis) to (I_i catalysis only) as the advantage of being covalent in each leg of the cycle. These ratios, indicated above the coloured arrows in Fig. 1b, are 1.73, 1.02 and 1.22 for $i = 1, 2$ and 3, respectively. Assuming these values are multiplicative, the cooperative benefit is about 2.2 for the entire cycle.

An impediment to truly hyperbolic growth for such a system[4] is the occasional formation of non-productive complexes (for example, **W–Y•Z**) through partially complementary base pairing (Fig. 1a). We can detect such complexes (Supplementary Fig. 6), but when they are minimized by pre-folding each RNA separately, the yield after 2 h increases by 25–50% (Supplementary Fig. 7). As shown by heat–cool regimes, reverse reactions that have the net effect of breaking down covalent ribozymes into fragments may also have a small role in preventing hyperbolic growth (Supplementary Fig. 8).

Cooperation versus selfishness

Next we tested whether a three-membered cooperative system has the potential to have higher fitness than purely autocatalytic systems when placed in direct competition (Fig. 2). To construct 'selfish' autocatalytic subsystems (S_i), we reverted the IGS–IGS target pairs within each subsystem so that they would match. To create S_1 we used $_{\mathrm{GUG}}\mathbf{W}_{\mathrm{CAU}}$ and ***h*•X•Y•Z**, to create S_2 we used $_{\mathrm{GAG}}\mathbf{W}\bullet\mathbf{X}_{\mathrm{CUU}}$ and ***h*•Y•Z**, and to create S_3 we used $_{\mathrm{GCG}}\mathbf{W}\bullet\mathbf{X}\bullet\mathbf{Y}_{\mathrm{CGU}}$ and ***h*•Z**. Each of these subsystems replicates well in isolation. Upon mixing of RNAs, we tracked selfish and cooperative ribozymes by the composition (matched or mismatched, respectively) of the **W**-containing fragments because these contain the IGS and hence the most crucial genetic element (Fig. 2a). When we compared the total yield of $S_1 + S_2 + S_3$ to that of $I_1 + I_2 + I_3$, the former out-performed the latter at all time points (that is, selfishness wins in isolation). One reason for this result is that there would be less time delay in initiating covalent synthesis in the all-selfish system. However, when we placed all six subsystems (12 RNAs: $I_1 + I_2 + I_3 + S_1 + S_2 + S_3$) in the same reaction, the relative yields at later times are reversed, and the growth of the enzymes resulting from the cooperative network now exceeds those from the selfish subsystems (that is, cooperation wins in competition). These results are independent of the exact RNA fragments we chose, as the same result can be seen in other systems with different IGS and IGS targets (see Supplementary Fig. 9). The yield reversal upon mixing happens because the selfish enzymes now participate in—and effectively expand—the cooperative network (Supplementary Fig. 10). This would be a mechanism for a network connectivity increase when the subsystems involved are competing for at least one shared resource, in this case the catalytic core (**Y–Z**), because all **W**-containing fragments can use the same 3′ fragments. Whereas selfish enzymes can also benefit from the network, the asymmetry in the proficiencies of the various IGS–IGS-target pairings creates potential for an asymmetry in the relative benefits of the various enzymes in the mixed environment. This feature would have been common in primordial genetic systems, allowing us to posit that cooperation could have been predisposed even in homogeneously mixed environments.

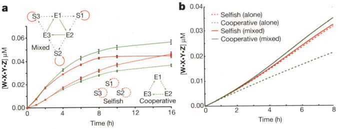


Figure 2 | Cooperative chemistry out-competes selfish chemistry when directly competed. **a**, Empirical results using cooperative (I_1, I_2 and I_3, that is, Fig. 1b) and selfish subsystems (S_1, S_2 and S_3, where IGS and IGS targets were changed to be matching in each subsystem). Yields of total **W•X•Y•Z** RNA tracked the concentrations of cooperative (mismatched) or selfish (matched) **W**-containing RNAs (0.05 μM initial concentrations) over time either when the cooperative (green) and selfish (red) sets of subsystems were incubated separately (dashed lines) or together in the same reaction mixture (solid lines; upper left inset). Data points are averages of three independent trials. Error bars show the standard error of the mean (s.e.m.), and the yields of the cooperative trials in the mixed experiment are significantly greater than those of the selfish trials at the 10- and 16-h time points ($P < 0.05$ by *t*-tests using Sidák's correction for multiple *a posteriori* comparisons). **b**, Simulation of growth dynamics using a toy model of the network of cooperation and selfish interactions (see Supplementary Information). Cooperative enzymes fare better in competition than do selfish enzymes, as demonstrated empirically in panel **a**.

Modelling

Empirical systems such as the one described above are subject to the particularities of chemical and methodological idiosyncrasies, so we sought to generalize these results by constructing mathematical models that show that under a certain set of parameters, the laboratory results should indeed be possible. First we constructed an ordinary differential equation (ODE) model for the three-membered network shown in Fig. 1b. We tracked the yield of each of the three E_i ribozymes separately—using three identical replicates from the same initial reaction mixture—by taking aliquots every 30 min for 16 h (Fig. 1d). We used standard optimization techniques to find the rate constants of all the possible reactions in Fig. 1b that produced trajectories in the ODE system closest to the observed data (Supplementary Information). We used these estimated rate constants to construct a second ODE model that would mimic the cooperative growth of the three subsystems. In general, the non-covalent versions of the ribozymes form relatively tight complexes, with K_d values in the low nanomolar range. When we built cooperative behaviour into the model by relying on differential equations of type $dE_j/dt = k_{ij}[I_j][E_i]$, the experimental data were fit very well in all three subsystems (Supplementary Fig. 11). When we removed direct catalysis from the model and inserted only autocatalysis instead, the quality of the fit decayed substantially such that the root mean squared error was 2.4-fold greater (Supplementary Fig. 12), confirming these results. These data support the contention that replication of the subsystems is indeed cooperative.

Next we constructed a toy model comparing the cooperative and selfish behaviours seen in Fig. 2a using the dynamical relationships that can exist among all enzymes (Fig. 2b). The 'selfish' enzymes perform some altruistic catalysis when alternative substrates become available. The empirical data display more striking yield differences than the model, perhaps because the time delays in bringing the results of the selfish catalytic events back to the selfish subsystems are exacerbated by physical processes such as diffusion. Again this result is general, at least within this network topology, and does not depend on the particular IGS–IGS-target pairings chosen. In essence, although the selfish replicators can parasitize the cooperators, the cooperative network benefits more by incorporating the selfish RNAs. Interestingly, the opposite is generally true in evolutionary dynamics: groups of cooperative individuals grow more quickly than groups of selfish individuals, but a group consisting of both types will eventually be dominated by the selfish[26]. One limitation to the experiment shown in Fig. 2a is that there is only a single iteration of selection. The RNAs used to seed the experiment limit its evolutionary potential; Supplementary Fig. 13 depicts joint genotype frequency changes over time. Experiments in a serial transfer format are needed to show the selection of one strategy over the other (see below), but we can use both our data and modelling to predict that cooperation would have been advantageous in simpler chemical systems that preceded organismal biology.

Randomization experiment

The system described in Fig. 1 is only one of a very large number of possibilities. To test the notion that cooperative networks of RNA could form spontaneously, we randomized the middle nucleotide of both the IGS (M) and its target triplet (N) in fragments of the ribozyme, generating both matched and mismatched partners within a population. We created three pools of randomized fragments containing the IGS on the 5′ end of the ribozyme: $_{\mathrm{GMG}}\mathbf{W}_{\mathrm{CNU}}$, $_{\mathrm{GMG}}\mathbf{W{\bullet}X}_{\mathrm{CNU}}$ and $_{\mathrm{GMG}}\mathbf{W{\bullet}X{\bullet}Y}_{\mathrm{CNU}}$, plus three fragments containing the catalytic core and the 3′ end of the ribozyme: **X•Y•Z**, **Y•Z** and **Z** (Fig. 3a). Fourfold variation in M and in N, combined with threefold variation in the junction (*j*) where recombination occurs (before **X**, **Y** or **Z**) leads to 48 genotypic possibilities (Fig. 3a). These assembled ribozymes can be distinguished by three variables: (1) the middle nucleotide of the IGS (M), (2) the location of the junction (x, y or z) and (3) the middle nucleotide of the target (N). We therefore denote each ribozyme with the three-letter code M*j*N, where *j* = x, y or z. Each of these ribozymes can be covalently assembled by any other ribozyme, itself covalently contiguous or not, provided that M in the catalyst is complementary to N in the substrate.

When we incubated equimolar amounts of these six RNA sets, all 48 possible full-length **W•X•Y•Z** *Azoarcus* ribozymes arose. The relative frequencies of the 48 possible full-length ribozymes recovered at each time point over an 8 h time course (Supplementary Table 2) show that, in accordance with the above and published data[25], recombination at the **Y**–**Z** junction is favoured, but no single genotype ever exceeded 13% of the total. Growth in the randomization experiment showed markedly greater yields (2–12-fold) than in our engineered three-membered system (Fig. 3b), indicating that far more productive interactions among RNA species are occurring in the former.

From approximately three million **W•X•Y•Z** genotypes sampled at each time point, distinct trends portray indirect evidence of a rapid succession from smaller to increasingly larger networks of cooperators (Fig. 3c, d). Genotypes that could easily propagate by selfish autocatalytic replication peak at or before the first time point at 30 min (Fig. 3c, dotted line with crosses). These are S_i genotypes (for example, those in Fig. 2) where M and N are complementary. A prime example is CyG, which could increase in number from the association of $_{\mathrm{GCG}}\mathbf{W{\bullet}X}_{\mathrm{CGU}}$ and **Y•Z** molecules, and this genotype rose in frequency from 4.8% to 7.2% between 30 min and 2 h. Out of the 48 possible product genotypes, twelve (25%) are of this type.

After peaking early, the frequencies of autocatalysts dropped below random expectation and then slowly climbed. Because of extremely large sample sizes, these deviations are highly significant (two-tailed *G*-tests of independence; $P \ll 0.001$). However, this later frequency increase may not be a consequence of autocatalysis per se, but of the incorporation of autocatalysts into higher-ordered networks, akin to the mechanism by which cooperative networks assimilate selfish replicators (Fig. 2 and Supplementary Fig. 10). Analyses of the frequencies of the product genotypes cannot reveal the identities of the catalysts that made them, and thus do not provide direct evidence of replicator cycles. Nevertheless, we examined whether networks of two or more distinct members could be increasing over time. Some pairs of genotypes can cooperate with each other to form two-membered cycles (for example, AxC + GzU), whereas others cannot (for

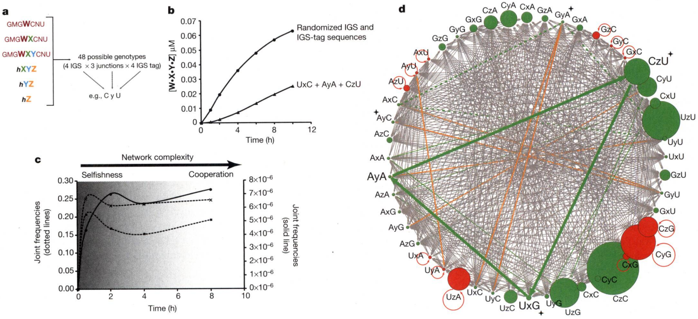


Figure 3 | The randomization experiment. a, Experimental design. The middle nucleotides of the IGS and the tags were randomized to create diverse RNA pools. A reaction of 300 pmol each (0.5 µM) of $_{\mathrm{GMG}}\mathbf{W}_{\mathrm{CNU}}$, $_{\mathrm{GMG}}\mathbf{W{\bullet}X}_{\mathrm{CNU}}$, $_{\mathrm{GMG}}\mathbf{W{\bullet}X{\bullet}Y}_{\mathrm{CNU}}$, **X•Y•Z**, **Y•Z** and **Z** was sampled at 0.5, 2, 4 and 8 h, and millions of recombined full-length **W•X•Y•Z** ribozymes were genotyped by nucleotide sequence analysis (Supplementary Table 2). **b**, Comparison of growth curves from fixed and randomized RNAs. Yields over time were compared for the simple three-membered cycle (filled triangles, UxG + AyA + CzU; the sum of the three curves in Fig. 1d) to that in the randomized format (filled circles, panel **a**) when both were performed at the same RNA pool concentrations (0.05 µM). **c**, Proposed succession from simple to complex networks using genotype frequency data from experiment in panel **a**. Simple autocatalytic cycles where M and N are complementary were directly tracked by the sum of such **W•X•Y•Z** molecules (dashed line with crosses; for example, AzU). Reciprocal two-membered cycles were tracked by the sum (×10, for ease of presentation) of the joint frequencies of all genotypes that can potentially participate in such cycles (dashed line with squares; for example, AxA + UxU). The rise of three-membered cycles can be seen from the sum (×10,000 for ease of presentation) of joint frequencies of three sets of genotypes: Fig. 1b and its two permutations by junction (solid line; UxG + AyA + CzU; UyG + AzA + CxU; UzG + AxA + CyU). See Supplementary Information for calculation of the joint frequencies. **d**, The potential network of RNA genotypes. Each node is one of the 48 possible M*j*N genotypes; size scales with relative frequency in the 8 h pool. Nodes are autocatalysts (red) or those that must replicate cooperatively (green). Grey arrows show all possible direct catalytic events; orange arrows show reciprocal two-membered cycles in which the frequencies of both members at least double between 30 min and 2 h; green arrows show key three-membered networks: thick green is the system studied in depth (Fig. 1b), thin green are permutations by junction, dotted green is AxC + GyA + UyU. Starred genotypes can participate in a four-membered network.

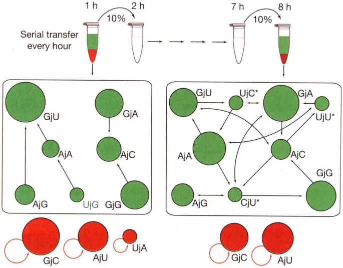


Figure 4 | The serial transfer experiment. The same RNA used to seed the randomization experiment (Fig. 3) was also subjected to a serial transfer protocol. For the first iteration, 50 pmol each of ${}_{\mathrm{GMG}}\mathbf{W}_{\mathrm{CNU}}$, ${}_{\mathrm{GMG}}\mathbf{W{\bullet}X}_{\mathrm{CNU}}$, ${}_{\mathrm{GMG}}\mathbf{W{\bullet}X{\bullet}Y}_{\mathrm{CNU}}$, **X•Y•Z**, **Y•Z** and **Z** were incubated in a 100 μl volume. After 1-h time points, 10% of the reaction mixture was transferred to a new tube containing 90% fresh RNA with a total volume of 100 μl. The population was sampled via 5′ RACE and RT–PCR to capture variation in all positions of any **W•X•Y•Z** molecules present in the population. The 1 and 8 h populations were cloned, and genotype frequencies were obtained by manual sequence analysis of 75 clones each (Supplementary Table 3). Any genotype present twice or more was included on this diagram (see text); size of the circles scales to relative frequencies within their respective populations. All possible catalytic interactions are shown with arrows among non-autocatalytic genotypes (green), with autocatalytic genotypes (red) not participating in the network. Grey genotype in the 1st iteration disappears. Genotypes with asterisks appear by the 8th iteration.

example, AxC + UzG). We noticed that the global joint frequencies of the members comprising all possible two-membered cycles peaked at 30 min, declined and recovered, although delayed with respect to the autocatalysts (Fig. 3c). Support for the succession from autocatalysts to these two-membered cycles is found in the frequencies of two possible partners for the autocatalysts GjC, which are CxG and CzG (autocatalysts themselves); the sum of these rose monotonically between 2–8 h (3.7% to 6.1%).

At roughly 2 h, a succession to three-membered cycles may have occurred. Although there are hundreds of such possible assemblages, the joint frequencies of the members of diverse ones requiring synthesis at all three junctions (such as UxG + AyA + CzU) jump at the 2 h mark (Fig. 3c, solid line). Many others peak then as well; the joint frequency of the AxC + GyA + UyU trio increases nearly 20-fold after the 30 min point. At 4 h and later the possibility of succession to even higher-ordered networks that subsume all simpler ones obfuscates individual trends. Visualization of all possible connections among genotypes underscores these conclusions (Fig. 3d). By 8 h the network is dominated by genotypes that can only be replicated via cooperation (green circles). In fact, the variance in the genotype frequencies drops monotonically over the course of the experiment, indicating that all genotypes increasingly participate in the network over time.

Serial transfer of the randomized population

The experiments depicted in Fig. 3 portray the dynamic changes that occur on a kinetic time scale as a batch of RNAs approaches equilibrium. In an actual prebiotic scenario, however, this effect would be iterated and perhaps magnified over several generations, as opposed to being an asymptotic value that results from mixing several RNAs in a single reaction vessel. To bring a stronger evolutionary flavour, we repeated the randomization experiment but in a serial transfer format. Starting with another aliquot of the exact same set of RNAs (that is, products from the same *in vitro* transcription), we carried a population through eight serial transfers, taking 10% of the population each hour into a fresh tube of fragments. In this manner the **W•X•Y•Z** molecules that spontaneously assemble are continually being fed with new fragments, such that selection will favour those molecules and networks that grow faster and persist over iterations. Given that the assembly that occurs each round can be strongly influenced by the actions of naive RNAs from the 90% fresh material, we opted to assay genotypic change by sampling only the most high-frequency genotypes: those present in an abundance greater than random chance (1/48). By manually sequencing the same number of genotypes (75) from transfers number 1 and 8 and enumerating those genotypes present more frequently than random expectation (2/75 > 1/48), we were able to observe the amalgamation of an RNA network over time (Fig. 4). At the 1 h time point, no closed network was possible and autocatalysts were relatively frequent (33%), but by 8 h a reflexively autocatalytic set was present in which every reaction is catalysed by at least one molecule involved in any of the reactions of the set[27]. This set included nine genotypes and fewer autocatalysts (25%), although the latter drop is not quite statistically significant (one-tailed *G*-test of independence; $P = 0.14$). Such expansion of the network to add additional genotypes is a more general case than the direct competition that we described in Fig. 2. As another indicator of the effect of serial transfer, the outcome of this experiment differed markedly from the batch assembly experiment (Fig. 3). After 8 h in the batch experiment the genotypes were dominated by pyrimidine-containing IGSs and targets (YzY; Fig. 3d). By contrast, the serial transfer experiment, although also reiterating the bias for the **Y–Z** junction, distinctly favoured IGS and target sequences containing purines (RzR; Fig. 4).

Fragmentation into four pieces

Lastly, we tested whether increased fragmentation of the RNA could provide additional complexity, and enhance the pre-biological relevance. We did this by breaking the molecule up into four pieces instead of two, creating four-piece versions of I_1, I_2 and I_3 analogously to Fig. 1a, b. When we mixed the resulting 12 RNAs together, we observed two interesting phenomena (Fig. 5). First, the growth curve was distinctly sigmoidal, indicating that when more fragments are involved, the cooperativity of the system becomes even more apparent. In the four-piece fragmentation, **W•X•Y•Z** ribozymes can be made via many pathways, including those in which more than one enzyme cooperates to construct the product: for example, an E_1 ribozyme could recombine the **W–X** junction, an E_2 ribozyme could recombine the **X–Y** junction, and an E_3 ribozyme could recombine the **Y–Z** junction. Second, analysis of the sequences of the product **W•X•Y•Z** ribozymes showed that such cooperation was common (Supplementary Fig. 14). In fact 85% of all ribozymes required help from enzymes from at least two subsystems (Fig. 5).

Discussion

Our results illustrate a scenario in which simple autocatalytic cycles form easily but are later supplanted by more complex cooperative networks that take advantage of the autocatalysts. Our system describes the short-term kinetic phenomena that provide the foundation for evolutionary behaviour[10] in the presence of sequence variation throughout the ribozymes analogous to those described as "prelife"[9]. Features of the system described here that would make it relevant to early evolution are that it is comprised solely of RNA (although other polymers could display cooperative behaviour[17,18]) and that the 3-nt IGS or IGS targets are essentially the tag sequences[28] that have been suggested as a means to form molecular coalitions that can partition genetic information in a homogeneous milieu. Closure of autocatalytic sets would have been facilitated by the cooperative aggregation of oligomers with related tags[21]. Subsequent expansion of cooperative networks as shown here is possible by invasion of the network by a new set with a distinct tag sequence, for example, moving from the

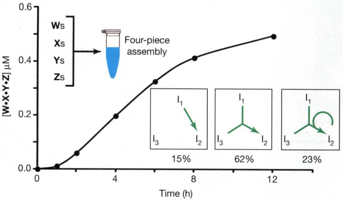


Figure 5 | Growth curve of a four-piece system. A more highly fragmented system based on that shown in Fig. 1b was created by breaking the ribozyme into four fragments for each I_i subsystem. The resulting 12 RNAs were co-incubated at 0.5 μM each, and samples were removed over time for both yield analysis (plot) and nucleotide sequence analysis (frequencies). The **W•X•Y•Z** RNAs can be assembled from a minimum of one, two or three IGS-bearing enzymes (examples shown with diagrams); the high frequencies of the latter two classes demonstrates the system's cooperativity.

three-membered cycle to a four-membered cycle such as by inclusion of a new IGS–IGS-target pair (Fig. 3d, starred genotypes), and then well beyond four members (Fig. 4). Longer-term evolutionary optimization would have required spatial heterogeneity[29] or compartmentalization[8,30] to provide lasting immunity against parasitic species or short autocatalytic cycles. Over time, a transition back to purely selfish replicators based on polymerization chemistry could proceed[7].

In our system, we show how RNA networks have the potential to arise spontaneously and to buffer informational decay. A key to the latter is the use of recombination for replication. Although allowing for some genotypic variability, it does not lead to the accumulation of deleterious mutations as does template-directed polymerization[31]. Highly interdependent networks of genetically related replicators as a means to circumvent the error catastrophe in nascent life have been proposed[11]. The three-membered cycle shown here resembles a hypercycle as envisioned previously[4,21,32], but without hyperbolic growth. We prefer to focus on the observation that the cycle can be derived from simpler cycles and has the potential to expand to more complex ones as evidence that RNA molecular coalitions can show spontaneous order-producing dynamics, which already has theoretical support[27]. Molecular ecological succession is a plausible model for a bridge between selfish replicators and cooperative systems.

METHODS SUMMARY

Experimental. Ribozyme assays or covalent self-assembly from oligomers were performed as described previously[24,25]. Briefly, RNA oligomers were incubated together in 100 mM $MgCl_2$ and 30 mM EPPS buffer (pH 7.5) for 5 min–16 h at 48 °C at a final concentration of 0.01–2.0 μM each. Visualization and quantification was possible via phosphorimaging when **W**-containing fragments were 5′-end-labelled with $\gamma[^{32}P]$ATP before use. For genotyping, $\sim$200-nt RNA was excised from a gel and subject to PCR with reverse transcription (RT–PCR) using **W**- and **Z**-specific primers. High-throughput sequence analysis on the Illumina platform was possible after 5′ RACE to capture the sequence variability in the IGS of assembled ribozymes. For manual sequence analysis, the PCR products were cloned into *Escherichia coli* and individual colonies were picked for colony PCR reactions. Resulting amplicons were either subjected to nucleotide-sequence analysis or restriction fragment length polymorphism (RFLP) analysis.

Modelling. The cooperative system was modelled as a set of six differential equations describing the concentrations over time of the six principal species (see Supplementary Information). These equations are derived from the detectable catalysis reactions (encompassing six direct-catalysis reactions and three cross-catalysis reactions). The experimental time series data from the full three-component system and from the two-component subsystems yielding detectable product were fit simultaneously to the model by standard optimization techniques.

Received 11 March 2011; accepted 6 September 2012.
Published online 17 October; corrected online 31 October 2012.

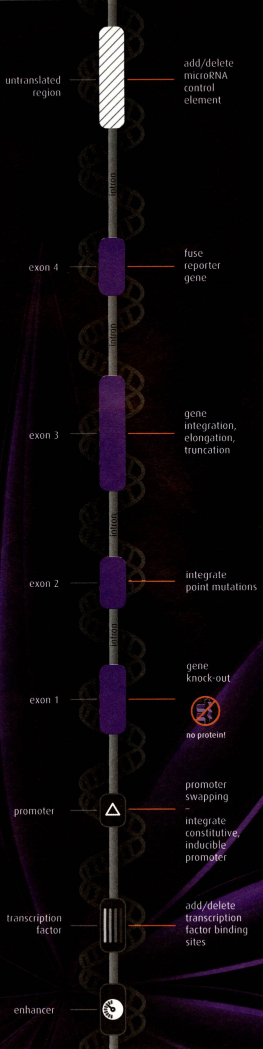
1. Joyce, G. F. RNA evolution and the origins of life. *Nature* **338,** 217–224 (1989).
2. Zaher, H. S. & Unrau, P. J. Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. *RNA* **13,** 1017–1026 (2007).
3. Wochner, A., Attwater, J., Coulson, A. & Holliger, P. Ribozyme-catalyzed transcription of an active ribozyme. *Science* **332,** 209–212 (2011).
4. Eigen, M. & Schuster, P. The hypercycle. A principle of natural self-organization. Part A: emergence of the hypercycle. *Naturwissenschaften* **64,** 541–565 (1977).
5. Kauffman, S. A. *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford Univ. Press, 1993).
6. Sievers, D. & von Kiedrowski, G. Self-replication of complementary nucleotide-based oligomers. *Nature* **369,** 221–224 (1994).
7. Levy, M. & Ellington, A. D. The descent of polymerization. *Nature Struct. Biol.* **8,** 580–582 (2001).
8. Szathmáry, E. The origin of replicators and reproducers. *Phil. Trans. Royal Soc. B* **361,** 1761–1776 (2006).
9. Nowak, M. A. & Ohtsuki, H. Preevolutionary dynamics and the origin of evolution. *Proc. Natl Acad. Sci. USA* **105,** 14924–14927 (2008).
10. Eigen, M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58,** 465–523 (1971).
11. Maynard Smith, J. Hypercycles and the origin of life. *Nature* **280,** 445–446 (1979).
12. Kun, Á., Santos, M. & Szathmáry, E. Real ribozymes suggest a relaxed error threshold. *Nature Genet.* **37,** 1008–1011 (2005).
13. Doudna, J. A. & Szostak, J. W. RNA-catalysed synthesis of complementary-strand RNA. *Nature* **339,** 519–522 (1989).
14. Joyce, G. F. Forty years of *in vitro* evolution. *Angew. Chem. Int. Ed.* **46,** 6420–6436 (2007).
15. Lincoln, T. A. & Joyce, G. F. Self-sustained replication of an RNA enzyme. *Science* **323,** 1229–1232 (2009).
16. Levy, M. & Ellington, A. D. Exponential growth by cross-catalytic cleavage of deoxyribozymogens. *Proc. Natl Acad. Sci. USA* **100,** 6416–6421 (2003).
17. Lee, D. H., Severin, K., Yokobayashi, Y. & Ghadiri, M. R. Emergence of symbiosis in peptide self-replication through a hypercyclic network. *Nature* **390,** 591–594 (1997).
18. Lee, D. H., Severin, K. & Ghadiri, M. R. Autocatalytic networks: the transition from molecular self-replication to molecular ecosystems. *Curr. Opin. Chem. Biol.* **1,** 491–496 (1997).
19. Voytek, S. B. & Joyce, G. F. Niche partitioning in the coevolution of 2 distinct RNA enzymes. *Proc. Natl Acad. Sci. USA* **106,** 7780–7785 (2009).
20. Kim, D.-E. & Joyce, G. F. Cross-catalytic replication of an RNA ligase ribozyme. *Chem. Biol.* **11,** 1505–1512 (2004).
21. Eigen, M. & Schuster, P. The hypercycle. A principle of natural self-organization. Part C: the realistic hypercycle. *Naturwissenschaften* **65,** 341–369 (1978).
22. Reinhold-Hurek, B. & Shub, D. A. Self-splicing introns in tRNA genes of widely divergent bacteria. *Nature* **357,** 173–176 (1992).
23. Hayden, E. J. & Lehman, N. Self-assembly of a group I intron from inactive oligonucleotide fragments. *Chem. Biol.* **13,** 909–918 (2006).
24. Hayden, E. J., von Kiedrowski, G. & Lehman, N. Systems chemistry on ribozyme self-construction: evidence for anabolic autocatalysis in a recombination network. *Angew. Chem. Int. Ed.* **47,** 8424–8428 (2008).
25. Draper, W. E., Hayden, E. J. & Lehman, N. Mechanisms of covalent self-assembly of the *Azoarcus* ribozyme from four fragment oligonucleotides. *Nucleic Acids Res.* **36,** 520–531 (2008).
26. Nowak, M. A. *Evolutionary Dynamics: Exploring the Equations of Life* (Harvard Univ. Press, 2006).
27. Hordijk, W. & Steel, M. Detecting autocatalytic, self-containing sets in chemical reaction systems. *J. Theor. Biol.* **227,** 451–461 (2004).
28. Weiner, A. M. & Maizels, N. 3′ terminal tRNA-like structures tag genomic RNA molecules for replication: implications for the origin of protein synthesis. *Proc. Natl Acad. Sci. USA* **84,** 7383–7387 (1987).
29. Boerlijst, M. C. & Hogeweg, P. Spiral wave structures in prebiotic evolution: hypercycles stable against parasites. *Physica D* **48,** 17–28 (1991).
30. Szathmáry, E. & Demeter, L. Group selection of early replicators and the origin of life. *J. Theor. Biol.* **128,** 463–486 (1987).
31. Lynch, M., Burger, R., Butcher, D. & Gabriel, W. The mutational meltdown in asexual populations. *J. Hered.* **84,** 339–344 (1993).
32. Eigen, M. & Schuster, P. The hypercycle. A principle of natural self-organization. Part B: the abstract hypercycle. *Naturwissenschaften* **65,** 7–41 (1978).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank A. Burton, R. Ghadiri, P. Higgs, B. Larson, K. Chacón and A. López García de Lomana for help during preparation of this manuscript. This work was supported by NASA grant NNX10AR15G to N.L., the Center for Life in Extreme Environments at Portland State University, and a fellowship from the Human Frontier Science Program Organization to R.X.-B.

Author Contributions N.L. and N.V. conceived the basic idea and conducted the experiments; E.J.H. and I.A.C. contributed to the evaluation of the results; I.A.C., M.L.M. and R.X.-B. constructed the mathematical models; N.L. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.L. (niles@pdx.edu).


untranslated region
add/delete microRNA control element
intron
exon 4
fuse reporter gene
intron
exon 3
gene integration, elongation, truncation
intron
exon 2
integrate point mutations
intron
exon 1
gene knock-out
no protein!
promoter
promoter swapping
–
integrate constitutive, inducible promoter
transcription factor
add/delete transcription factor binding sites
enhancer

TALEN™

Visit us at booth 1227 at ASHG
San Francisco
6-10 Nov 2012

edit any gene at any position in any genome

Custom TALEN™ and exclusive support for all your gene editing projects only available at Cellectis bioresearch

With Cellectis bioresearch Custom TALEN™ Services, all gene editing projects can now be achieved including gene knock-out, gene tagging, promoter swapping, nucleotide substitution, protein truncation, and many more.

With over fifteen years expertise in genome customization Cellectis bioresearch is the dedicated service provider in Custom TALEN™ design, production, validation and comprehensive project support. We know that the key to your success is both technology and know-how, and that's why our dedicated project support managers will provide you with dedicated expertise to achieve customized cells and organisms harboring your modification of choice.

No limit, just talent!

✱ for more information, please visit **cellectis-bioresearch.com**

cellecti
bioresearch

GENOME CUSTOMIZATI

doi:10.1038/nature11534

Distinctive space weathering on Vesta from regolith mixing processes

C. M. Pieters[1], E. Ammannito[2], D. T. Blewett[3], B. W. Denevi[3], M. C. De Sanctis[2], M. J. Gaffey[4], L. Le Corre[5], J.-Y. Li[6], S. Marchi[7], T. B. McCord[8], L. A. McFadden[9], D. W. Mittlefehldt[10], A. Nathues[5], E. Palmer[11], V. Reddy[5], C. A. Raymond[12] & C. T. Russell[13]

The surface of the asteroid Vesta has prominent near-infrared absorption bands characteristic of a range of pyroxenes, confirming a direct link to the basaltic howardite–eucrite–diogenite class of meteorites[1,2,3]. Processes active in the space environment produce 'space weathering' products that substantially weaken or mask such diagnostic absorption on airless bodies observed elsewhere[4,5], and it has long been a mystery why Vesta's absorption bands are so strong. Analyses of soil samples from both the Moon[6] and the asteroid Itokawa[7] determined that nanophase metallic particles (commonly nanophase iron) accumulate on the rims of regolith grains with time, accounting for an observed optical degradation. These nanophase particles, believed to be related to solar wind and micrometeoroid bombardment processes, leave unique spectroscopic signatures that can be measured remotely[8–10] but require sufficient spatial resolution to discern the geologic context and history of the surface, which has not been achieved for Vesta until now. Here we report that Vesta shows its own form of space weathering, which is quite different from that of other airless bodies visited. No evidence is detected on Vesta for accumulation of lunar-like nanophase iron on regolith particles, even though distinct material exposed at several fresh craters becomes gradually masked and fades into the background as the craters age. Instead, spectroscopic data reveal that on Vesta a locally homogenized upper regolith is generated with time through small-scale mixing of diverse surface components.

When well-developed (weathered, or 'mature') soils of the Moon are compared with freshly exposed material[4,11], they are found to be darker, to have weaker absorption bands and to exhibit a red-sloped continuum to a wavelength of 2.6 μm (reflectance increasing towards longer wavelength). Through decades of research, it has been recognized in the laboratory[6,10], experimentally[8] and with physical modelling[10] that the nanometre-scale particles of metallic iron (called nanophase iron, or npFe^0) that accumulates on the surface of soil grains produces all three of the observed optical properties of lunar soils. These properties have been used to evaluate space weathering on other airless bodies, but the results have been contentious, largely as a result of incomplete information or necessary simplifying assumptions. Because Vesta is bright and is found to have strong absorption bands in telescopic data, comparable to those of howardite–eucrite–diogenite (HED) meteorites[3,12,13], it was originally argued that space weathering does not occur to any significant extent on asteroids[14]. However, one of the largest groups of asteroids (S type) was found to have a mineralogy (low-calcium pyroxene and olivine) comparable to that of the most common class of meteorites (ordinary chondrites), but the asteroids also have a red-sloped continuum and much weaker diagnostic absorption bands than their possible meteorite counterparts. This required either that an asteroid–meteorite link be made between some S-type asteroids and ordinary chondrites by space weathering processes on the asteroid or, if no space weathering occurs, that alternative compositional interpretations be found to account for the measured optical properties[15,16].

This debate over space weathering produced two types of compositional interpretation of S-type asteroids, which implied dramatically different models of thermal evolution for the inner Solar System (one requiring melting and differentiation; the other allowing a preponderance of unprocessed chondritic materials). The resulting impasse began to be resolved when two S-type asteroids were visited by spacecraft. Data from the NEAR-Shoemaker spacecraft provided strong evidence that the S-type asteroid Eros was indeed chondritic[17] and space weathered[18]. Decisive new information came from the recent return of many grains of samples from the surface of the near-Earth S-type asteroid Itokawa by the Hayabusa mission. The samples conclusively showed not only that Itokawa is of LL chondritic composition[19] as predicted from remote sensing[20], but also that half of the grains studied contained thin rims of nanophase iron- and sulphur-rich particles[7], accounting for the previously observed, weak, lunar-like space weathering optical effects[21]. Chondrite meteorites are commonly rich in FeS; thus, the sulphur component of the Itokawa nanophase particles is particularly noteworthy because the lunar nanophase particles are dominated largely by Fe^0 alone.

Vesta is found to be both different from and similar to these other airless bodies and provides key information on processes active in the asteroid belt. Data from the Dawn mission used here are from an imaging spectrometer, the Visible and Infrared Spectrometer[1] (VIR), and a multispectral camera with bands across the shorter wavelengths, the Framing Camera[2]. The VIR data were acquired during the survey phase of the mission from an altitude of ~2,700 km and at a resolution of 690 m per pixel, and the Framing Camera data were acquired during the high-altitude mapping orbit (HAMO) phase from an altitude of ~680 km and at a resolution of 64 m per pixel.

As on the Moon, material freshly excavated by craters on Vesta is found to be quite distinct from surrounding soils, and fresh craters of all sizes commonly have a local disturbed area or system of rays radiating from the crater across nearby terrain. Canuleia, a typical fresh crater ~11 km in diameter, shows this diversity (Fig. 1). Several manifestations of bright and dark materials are observed on a variety of scales on Vesta[22,23], and small, morphologically fresh craters (<1 km in diameter) can be surrounded by either bright or dark ray material (Fig. 2). Vesta is also seen to have an extensive impact-generated regolith and widespread mass wasting[24] such as can be observed for the two large craters in Fig. 2. Exposures of distinct bright and dark materials hundreds of metres in extent are seen along the walls of several craters even though the craters no longer have prominent rays. Because craters with softened and slightly more subdued morphology than the freshest craters do not have bright (or dark) ray systems, processes must exist on Vesta that alter surface soil and effectively erase such features over time.

---

[1]Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA. [2]Istituto di Astrofisica e Planetologia Spaziali, INAF, ARTOV, 00133 Rome, Italy. [3]Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland 20723, USA. [4]Department of Space Studies, University of North Dakota, North Dakota 58202, USA. [5]Max-Planck Institute for Solar System Research, 37191 Katlenburg-Lindau, Germany. [6]Department of Astronomy, University of Maryland, College Park, Maryland 20742, USA. [7]NASA Lunar Science Institute, Boulder, Colorado 80302, USA. [8]Bear Fight Institute, 22 Fiddler's Road, Box 667, Winthrop, Washington 98862, USA. [9]NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA. [10]NASA Johnson Space Center, Houston, Texas 77058, USA. [11]Planetary Science Institute, 1700 East Fort Lowell Road, Suite 106, Tucson, Arizona 85719, USA. [12]Jet Propulsion Laboratory, Caltech, Pasadena, California 91109, USA. [13]University of California, Los Angeles, California 90095, USA.

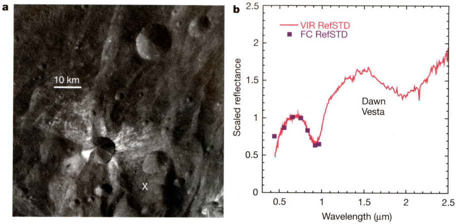


Figure 1 | The optical properties across a typical area on Vesta as measured by Dawn. a, Canuleia, a fresh crater (11.6 km in diameter, at 34° S, 295° E in the Dawn coordinate system[31]), has prominent bright rays. Most but not all rays are brighter than the surroundings. A crater of similar size with more subdued features (older) and no ray system is seen to the northeast. The white 'X' indicates a background reference region of soil used as a standard for comparison (RefSTD). Dawn Framing Camera HAMO image: FC21B0010859_11293124257F1A. **b,** Visible/near-infrared spectra of soil background reference region indicated in **a**, as acquired by the Dawn Framing Camera (FC) and the VIR. The Framing Camera obtained colour data for seven spectral bands[2] (0.44–0.98 μm) including half of a ferrous absorption, due to pyroxene, near 1 μm. As an imaging spectrometer, the VIR acquired images at lower spatial resolution but at high spectral resolution containing 864 spectral channels[1] (0.25–5.1 μm), which allow full assessment of the two ferrous absorptions of pyroxene near 1 and 2 μm. A colour composite of this area at higher spatial resolution, derived from lower-altitude Framing Camera data, can be found in Supplementary Fig. 1.

Spectra acquired by the VIR and Framing Camera are compared for the same large (~10-km) region of background soil unaffected by rays from Canuleia (Fig. 1) and are found to be largely in agreement. This background material is representative of average Vesta soil, and its spectrum has the prominent diagnostic absorption features of pyroxene that are characteristic of Vesta's basaltic composition[1–3]. The location of five areas of soil recently affected by the impacts that created Canuleia and Sossia craters are shown in Fig. 3. Spectra for these fresh (unweathered) areas have similar pyroxene features (Fig. 4). The bright and dark ray patterns are readily seen in photometrically corrected data (Supplementary Fig. 3). VIR spectra of these fresh areas and background soil are compared with analogous high-resolution spectra acquired by the Moon Mineralogy Mapper[25] for a fresh crater in basaltic lunar terrain (Fig. 4).

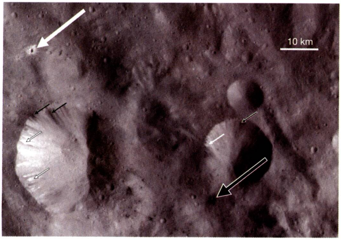


Figure 2 | Diverse material exposed at fresh craters in a well-developed regolith on Vesta. Examples of bright and dark rays surrounding small (<1-km), fresh craters (large white and black arrows) are observed along with discrete bright and dark material within the steep walls of older craters without rays (small arrows). Slump material dominates one wall of each of the two larger craters, illustrating the mobility of the regolith. The crater on the left is Helena (~20 km in diameter, at 41° S, 123° E in the Dawn coordinate system[31]). A photometrically corrected image containing this region can be found in Supplementary Fig. 2.

At the present spatial resolution of the Dawn data, all Vesta spectra are dominated by prominent absorptions due to different types of pyroxene[1]. On a local scale, where compositional variations are relatively small, it can be seen that brighter areas typically have stronger pyroxene absorption bands, and there is an overall correlation of brightness with the strength of the absorption bands[22]. Only one area (no. 2; see Fig. 3) has spectral variations suggestive of the presence of a minor component of different mineral constituents (best seen in Fig. 4e, f). The spectra for the lunar fresh crater and background (mature) soils in Fig. 4 illustrate the three optical properties of space weathering linked to npFe^0 seen on the Moon: soils are darker, have weaker absorption bands and exhibit a red-sloped near-infrared continuum (0.7–1.5 μm). Although a correlation between brightness and band strength is seen for both Vesta and the Moon, there is an important difference between the two bodies in terms of the relationship between freshly exposed and background material: areas on Vesta do not exhibit the notable change in continuum slope across the near-infrared that is observed for background soils of the Moon.

The observed lack of a consistent and systematic change in the near-infrared continuum slope on Vesta indicates that grain coating of opaque, nanophase iron particles comparable to that produced during space weathering of the Moon and smaller asteroids seems not to be present in any substantial amount on Vesta. This conclusion is consistent with detailed analyses of lunar and HED regolith breccias that observed a sparsity of lunar-like nanophase iron rims in HED meteorites[26]. Furthermore, the relative spectra in Fig. 4e, f indicate that across the visible part of the spectrum (0.5–0.7 μm) these freshly exposed dark areas (for example no. 5) are relatively 'bluer' than background soil, and that freshly exposed bright areas (for example nos 1 and 3) are relatively 'redder' than background soil. This is opposite to the predicted effect of the presence of npFe^0 grain coatings[4,5,10], but is more consistent with the presence of a finely dispersed, micrometre-size opaque phase that darkens and lowers the spectral contrast of a mafic-rich host[27].

Because it is now recognized that the cratered surface of Vesta has not been resurfaced recently and that the giant Rheasilvia basin at the asteroid's south pole is no younger than ~1 Gyr (refs 28, 29), Vesta's

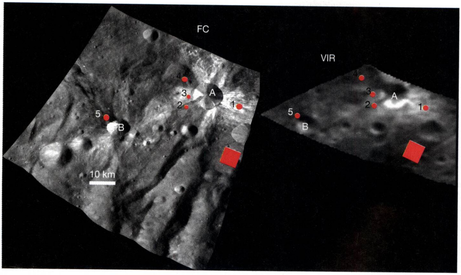


Figure 3 | The Canuleia region mapped at different spatial resolutions by Dawn's two optical instruments. Framing Camera and VIR images acquired during the Dawn HAMO and survey phases, respectively, contain the two morphologically fresh craters Canuleia (A) and Sossia (B). Spectra for five areas (indicated with small red circles, numbered from brightest to darkest in VIR data) were extracted as examples to highlight the diversity seen for freshly exposed material along bright (nos 1 and 3) and dark (no. 5) rays of fresh craters. The large red square is the area used as a background reference standard for both VIR and Framing Camera data on this region. A photometrically corrected image containing Canuleia and Sossia can be found in Supplementary Fig. 3 and clarifies the scale and extent of bright and dark variations across the region.

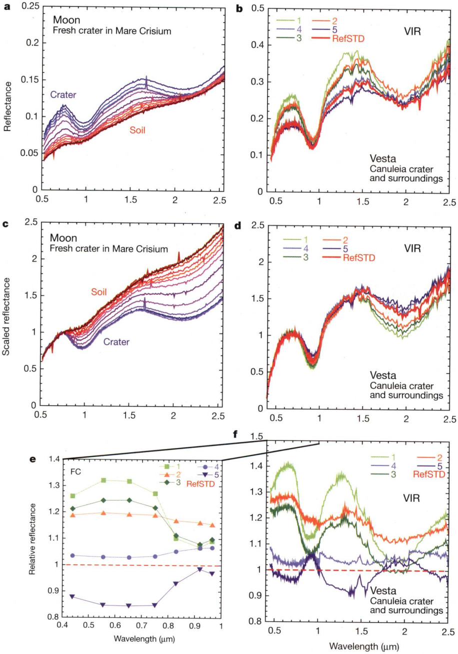


Figure 4 | Spectroscopic comparisons of Vesta soils with similar materials on the Moon. Spectra of the five areas of freshly exposed material and background regolith at Canuleia and Sossia identified in Fig. 3 are compared with spectra of comparable material on the Moon (presented as a traverse from an unnamed fresh crater towards background soils). **a**, **b**, Reflectance spectra illustrating brightness variations between crater materials and background soil (red spectra). **c**, **d**, Reflectance spectra scaled to 1 at 0.75 μm to allow comparison of spectral features. Lunar soils have a steeper continuum from 0.75 to 1.50 μm compared with the fresh crater, whereas there is no continuum change for Vesta. **e**, **f**, Reflectance spectra of Canuleia and Sossia areas in **b** divided by the background reflectance (red spectrum). These relative reflectance spectra maintain measured brightness differences, but capture subtle spectral variations and illustrate the excellent agreement between instruments despite the different illumination conditions of the measurements. See Supplementary Information for Framing Camera spectral traverses from Canuleia across background soils in a form similar to **e** (Supplementary Fig. 4); high-spatial-resolution images of the fresh lunar crater that provides the spectra in **a** and **c** (Supplementary Fig. 5); and spectra of howardite meteorites, prepared similarly to **b** and **f** (Supplementary Fig. 6).

surface is dominated by at least 1 Gyr of regolith evolution. Although distinctive material freshly exposed on the surface of Vesta by recent craters is indeed shown to blend into surrounding background regolith with time, space weathering on Vesta is quite different from that on other airless bodies. Regolith evolution on Vesta involves large- and small-scale impact events that produce particulate material and redistribute components. Compared with impact velocities of $\sim$15 km s^{-1} for bodies colliding at a distance from the Sun of one astronomical unit (where the Earth and Moon reside), the low average velocity[30] of $\sim$5 km s^{-1} expected for Vesta's location in the main asteroid belt suggest that mechanical brecciation dominates over melting and vaporization. Vesta's regolith does not accumulate detectible nanophase opaque particles on rims of grains. Instead, physical regolith processes on Vesta are sufficiently robust to alter optical properties by producing a locally well-mixed surficial regolith derived from observed host lithologies that include diverse bright and dark components[22,23]. The surficial regolith is mixed locally by smaller impacts that continually stir the regolith, and the process is enhanced by high regolith mobility driven by Vesta's relatively steep topography[24] and local gravity gradients. The new information about the surface of Vesta forces our concept of space weathering to go beyond one focused solely on solar wind and micrometeoroid interactions. We must include regolith mobility and fine-scale mixing as part of the complex assortment of global processes acting on an airless planetary surface over time.

Received 17 June; accepted 20 August 2012.

1. De Sanctis, M. C. *et al.* Spectroscopic characterization of mineralogy and its diversity across Vesta. *Science* **336,** 697–700 (2012).
2. Reddy, V. *et al.* Color and albedo heterogeneity of Vesta from Dawn. *Science* **336,** 700–704 (2012).
3. McCord, T. B., Adams, J. B. & Johnson, T. V. Asteroid Vesta: spectral reflectivity and compositional implications. *Science* **168,** 1445–1447 (1970).
4. Pieters, C. M. *et al.* Space weathering on airless bodies: resolving a mystery with lunar samples. *Meteorit. Planet. Sci.* **35,** 1101–1107 (2000).
5. Hapke, B. Space weathering from Mercury to the asteroid belt. *J. Geophys. Res.* **106,** 10039–10073 (2001).
6. Keller, L. P. & McKay, D. S. The nature and origin of rims on lunar soil grains. *Geochim. Cosmochim. Acta* **61,** 2331–2341 (1997).
7. Noguchi, T. *et al.* Incipient space weathering observed on the surface of Itokawa dust particles. *Science* **333,** 1121–1125 (2011).
8. Sasaki, S., Nakamura, K., Hamabe, Y., Kurahashi, E. & Hiroi, H. Production of iron nanoparticles by laser irradiation in a simulation of lunar-like space weathering. *Nature* **410,** 555–557 (2001).
9. Marchi, S., Brunetto, R., Magrin, S., Lazzarin, M. & Gandolfi, D. Space weathering of near-Earth and main belt silicate-rich asteroids: observations and ion irradiation experiments. *Astron. Astrophys.* **443,** 769–775 (2005).
10. Noble, S. K., Pieters, C. M. & Keller, L. P. An experimental approach to understanding the optical effects of space weathering. *Icarus* **192,** 629–642 (2007).
11. Adams, J. B. & McCord, T. B. Alteration of lunar optical properties: age and composition effects. *Science* **171,** 567–571 (1971).
12. Hiroi, T., Pieters, C. & Takeda, H. Grain size of the surface regolith of asteroid 4 Vesta estimated from its reflectance spectrum in comparison with HED meteorites. *Meteoritics* **29,** 394–396 (1994).
13. Pieters, C. M. *et al.* in *Asteroids, Comets, and Meteors* 273–288 (Cambridge Univ. Press, 2006).
14. Matson, D. L., Johnson, T. V. & Veeder, G. J. Soil maturity and planetary regoliths: the Moon, Mercury, and the asteroids. *Proc. Lunar Sci. Conf.* **8,** 625–627 (1977).
15. Clark, B. E., Hapke, B., Pieters, C. & Britt, D. in *Asteroids III* (eds Bottke, W. F., Cellino, A., Paolicchi, P. & Binzel, R. P.) 585–599 (Univ. Arizona Press, 2002).
16. Chapman, C. R. Space weathering of asteroid surfaces. *Annu. Rev. Earth Planet. Sci.* **32,** 539–567 (2004).
17. Trombka, J. *et al.* The elemental composition of asteroid 433 Eros: results of the NEAR-Shoemaker X-ray spectrometer. *Science* **289,** 2101–2105 (2000).
18. Clark, B. E. *et al.* Space weathering on Eros: constraints from albedo and spectral measurements of Psyche crater. *Meteorit. Planet. Sci.* **36,** 1617–1637 (2001).
19. Nakamura, T. *et al.* Itokawa dust particles: a direct link between S-type asteroids and ordinary chondrites. *Science* **333,** 1113–1116 (2011).
20. Binzel, R. P. *et al.* MUSES-C target asteroid (25143) 1998 SF36: a reddened ordinary chondrite. *Meteorit. Planet. Sci.* **36,** 1167–1172 (2001).
21. Hiroi, T. *et al.* Developing space weathering on the asteroid 25143 Itokawa. *Nature* **443,** 56–58 (2006).
22. McCord, T. *et al.* Dark material on Vesta from the infall of carbonaceous volatile-rich material. *Nature* http://dx.doi.org/10.1038/nature11561 (this issue).
23. Li, J.-Y. *et al.* Investigating the origin of bright materials on Vesta: synthesis, conclusions, and implications. *Proc. Lunar Planet. Sci. Conf.* **43,** abstr. 2381 (2012).
24. Jaumann, R. *et al.* Vesta's shape and morphology. *Science* **336,** 687–690 (2012).
25. Pieters, C. M. *et al.* The Moon Mineralogy Mapper (M^3) on Chandrayaan-1. *Curr. Sci.* **96,** 500–505 (2009).
26. Noble, S. K., Keller, L. P., &. Pieters, C. M. Evidence of space weathering in regolith breccias II: asteroidal regolith breccias. *Meteorit. Planet. Sci.* **45,** 2007–2015 (2010).
27. Britt, D. & Pieters, C. M. Darkening in black and gas-rich ordinary chondrite meteorites: the spectral effects of opaque morphology and distribution. *Geochim. Cosmochim. Acta* **58,** 3905–3919 (1994).
28. Marchi, S. *et al.* The violent collisional history of asteroid 4 Vesta. *Science* **336,** 690–694 (2012).
29. Schenk, P. *et al.* The geologically recent giant impact basin at Vesta's south pole. *Science* **336,** 694–697 (2012).
30. Bottke, W. F. & Nolan, M. C. Greenberg, R. & Kolvoord, R. A. Velocity distributions among colliding asteroids. *Icarus* **107,** 255–268 (1994).
31. Russell, C. T. *et al.* Dawn at Vesta: testing the protoplanetary paradigm. *Science* **336,** 684–686 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge the Dawn Instrument, Flight and Operations teams for the successful development, cruise, orbital insertion and operations of the Dawn spacecraft at Vesta. US team members are supported by the NASA Discovery Program through contract NNM05AA86C to the University of California, Los Angeles and by the NASA Dawn participating scientist programme.

Author Contributions C.M.P., D.T.B., M.C.D.S., S.M., L.A.M., D.W.M., V.R. and C.T.R. contributed to writing and improving the manuscript. E.A., L.L.C. and A.N. provided calibrated Dawn data. E.P. provided data searches. All authors contributed to discussion of the results.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.M.P. (carle_pieters@brown.edu).

LETTER

doi:10.1038/nature11561

Dark material on Vesta from the infall of carbonaceous volatile-rich material

T. B. McCord[1], J.-Y. Li[2], J.-P. Combe[1], H. Y. McSween[3], R. Jaumann[4], V. Reddy[5,6], F. Tosi[7], D. A. Williams[8], D. T. Blewett[9], D. Turrini[7], E. Palomba[7], C. M. Pieters[10], M. C. De Sanctis[7], E. Ammannito[7], M. T. Capria[7], L. Le Corre[5], A. Longobardo[7], A. Nathues[5], D. W. Mittlefehldt[11], S. E. Schröder[5], H. Hiesinger[12], A. W. Beck[13], F. Capaccioni[7], U. Carsenty[4], H. U. Keller[14], B. W. Denevi[9], J. M. Sunshine[15], C. A. Raymond[16] & C. T. Russell[17]

Localized dark and bright materials, often with extremely different albedos, were recently found on Vesta's surface[1,2]. The range of albedos is among the largest observed on Solar System rocky bodies. These dark materials, often associated with craters, appear in ejecta and crater walls, and their pyroxene absorption strengths are correlated with material brightness. It was tentatively suggested that the dark material on Vesta could be either exogenic, from carbon-rich, low-velocity impactors, or endogenic, from freshly exposed mafic material or impact melt, created or exposed by impacts. Here we report Vesta spectra and images and use them to derive and interpret the properties of the 'pure' dark and bright materials. We argue that the dark material is mainly from infall of hydrated carbonaceous material (like that found in a major class of meteorites and some comet surfaces[3–5]), whereas the bright material is the uncontaminated indigenous Vesta basaltic soil. Dark material from low-albedo impactors is diffused over time through the Vestan regolith by impact mixing, creating broader, diffuse darker regions and finally Vesta's background surface material. This is consistent with howardite–eucrite–diogenite meteorites coming from Vesta.

Vesta has a mean diameter[6] of 525 km and is the second-most massive object in the main asteroid belt of our Solar System, smaller than Ceres and similar to Pallas. These three bodies form a separate class of intact objects in the asteroid belt that have experienced planetary processes[7], such as thermal evolution[8] powered by short-lived radionuclides incorporated at the time of accretion. This process in general results in mineralogical alteration due to heating, and differentiation, with denser materials sinking towards the centre. In contrast, most other main-belt asteroids seem to be pieces of collisionally disrupted objects. Although subdued albedo differences on global[9] and broadly regional[10] scales were known to exist from telescopic observations, localized and intense dark and bright occurrences were not anticipated (see Supplementary Information).

We used photometrically uniform near-global albedo spectral image mosaics constructed from the Dawn[11] Framing Camera and

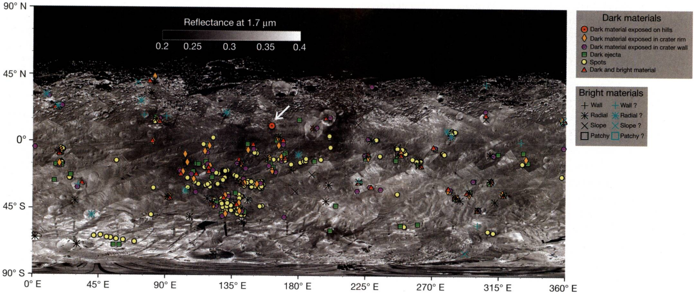


Figure 1 | Locations of dark and bright material. Mapped occurrences of local dark- and bright-material locations (points) are shown here plotted on a 1.7-μm albedo map derived from VIR images. The albedo mosaic also exhibits broad low-albedo regions, especially between about 70° and 220° longitude, near which the localized dark-material points tend to cluster, suggesting a causal relationship. Bright- and dark-material local examples tend not to be uniformly distributed or correlated with each other. For this base map, the entire VIR infrared data set from late Approach, Survey and High-Altitude Mapping orbits was converted into reflectance and mosaicked.

---

[1]Bear Fight Institute, 22 Fiddler's Road, Box 667, Winthrop, Washington 98862, USA. [2]Planetary Science Institute, Tucson, Arizona 85719, USA. [3]University of Tennessee, Knoxville, Tennessee 37996, USA. [4]DLR, Institute of Planetary Research, Berlin, 80302, Germany. [5]Max Planck Institute for Solar System Research, Max-Planck-Strasse 2, 37191 Katlenburg, Germany. [6]University of North Dakota, Grand Forks, North Dakota 58202, USA. [7]Istituto Nazionale di Astrofisica (IASF), Via Fosso del Cavaliere 100, Rome 00133, Italy. [8]ASU, Tempe, Arizona 85287, USA. [9]Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland 20723, USA. [10]Brown University, Providence, Rhode Island 02912, USA. [11]Astromaterials Research Office, NASA Johnson Space Center, Houston, Texas 77058, USA. [12]Institut für Planetologie, Westfälische Wilhelms-Universität Münster, Schlossplatz 2, 48149 Munster, Germany. [13]Department of Mineral Sciences, Smithsonian National Museum of Natural History, Washington DC 20024, USA. [14]Institut für Geophysik und extraterrestrische Physik, Mendelssohnstrasse 3, 38106 Braunschweig, Germany. [15]Department of Astronomy, University of Maryland, College Park, Maryland 20742, USA. [16]California Institute of Technology, Jet Propulsion Laboratory, Pasadena, California 91109, USA. [17]University of California, Los Angeles, California 90095, USA.

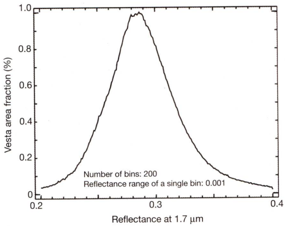


Figure 2 | Frequency distribution of albedos. The global distribution of albedos at 1.7 µm, derived from the infrared global base map (Fig. 1) at a spatial resolution of about 1 km. The curve represents surface area fraction (not number of pixels). At higher resolutions this distribution would develop smaller peaks at each extreme of brightness, representing individual dark- and bright-material locations. At this resolution, the broad global distribution represents mixing between dark and bright materials.

the visible and infrared mapping spectrometer (VIR) to demonstrate that dark and bright materials are distributed non-uniformly and are uncorrelated (as seen in Fig. 1), suggesting different origins. Unlike bright material, some dark material also occurs over larger areas with more diffuse boundaries.

Investigation of the 1.7-µm albedo frequency distribution at the spatial scale of 1 km per pixel surprisingly showed it to be unimodal, lacking any special albedo classes (Fig. 2). The bright material appears to be distinct from dark material on a local scale[1,2], but globally at this resolution these albedo classes form a continuum distribution with a peak at the global average albedo. This suggests a mixing of dark and bright materials to produce the range of Vesta surface materials.

We compared VIR infrared spectra for dark and bright material (Fig. 3). The pyroxene signature clearly dominates Vesta's spectra globally[8] and in detail[2,12], a product of Vesta's igneous past[8]. Dark material has weaker apparent absorptions[2]. We then applied a Multiple-Endmember Linear Spectral Unmixing Model[13] to derive the spectrum of each of Vesta's surface materials within each VIR pixel, using the fewest possible spectral endmembers. We found that the VIR spectra could be modelled using the weighted sum of only two spectral endmembers, called 'bright' and 'dark' materials in Fig. 3. The modelled dark-material endmember spectrum shows no absorption bands, with a reddish slope that flattens towards longer wavelengths. The modelled bright-material endmember spectrum shows the classic, strong pyroxene 2-µm band and the expected pyroxene continuum shape[8,12]. Our conclusion is that, for the most part, Vesta's surface material can be thought of as having two spectral components—the dark and bright endmember spectra—in different proportions. This is consistent with the mixing-process hypothesis and suggests that the dark spectral component is the agent diluting the pyroxene spectral signature. The identity of the bright endmember is probably the intrinsic Vesta basaltic soil, rich in unaltered, crystalline pyroxenes. This relationship between bright and dark materials is further supported by the strong correlation of the 1-µm pyroxene absorption band strength with albedo from analysis of the Framing Camera colour data[2] and shown here (Fig. 4). We note that the dark-material spectrum is very similar to that of carbonaceous chondrite material, such as is found in a major class of meteorites (the carbonaceous chondrites).

Regions of Vesta where the residuals from the spectral mixing analysis are largest (but still small) have some of the strongest pyroxene signatures. These represent the best opportunity to study intrinsic Vesta material, for example, at some apparent impact structures that may be sampling ejecta from the Rheasilvia basin near the south pole.

Suggestions for the origin of discrete areas of dark material include indigenous sources (such as opaque-rich lava flows and impact melts) as well as contamination from exogenous material (delivered from foreign, impacting bodies). A search of Dawn VIR spectra for OH spectral features near 3 µm was made[14,15], prompted by the discovery of OH and H_2O in the lunar surface. A 2.8-µm absorption was found[16] and its analysis indicates[17] that the dark material appears to be relatively enriched in OH (Fig. 5). Carbonaceous chondrite meteorite material often contains 10–20% OH-bearing hydrated minerals, whereas none of the other suggested dark-material sources contain

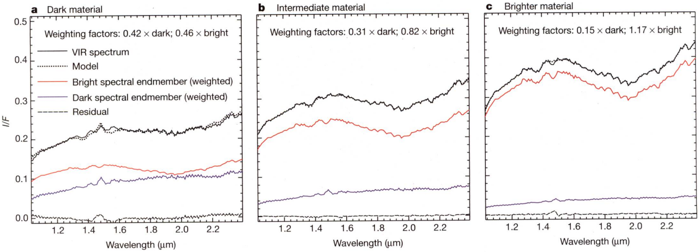


Figure 3 | Reflectance spectra of dark and bright materials. $I/F = \pi R/(d^2 F)$, where R is the spectral radiance of the target's surface in units of W m^{-2} sr^{-1} µm^{-1}, F is the spectral irradiance or solar flux in units of W m^{-2} µm^{-1} and d is the distance between the Sun and the target. Reflectance spectra for a representative VIR scene (366894613) illustrate Vesta spectra modelling results for three different types of units near the region indicated by the white arrow in Fig. 1. **a**, The VIR dark-material unit spectrum (solid black line) is modelled using the two weighted endmember spectra (red and blue spectra). The model result is the dotted black line overlying the VIR spectrum; the resulting residual is shown as a dashed black line. The residual shows no distinctive features and overall is near the noise level of the VIR data, suggesting successful modelling. (The feature near 1.44 µm is a calibration artefact.) **b**, **c**, Similar examples for intermediate and brighter materials nearby. The weighting factors are given on the plots and the mixture spectrum is calculated as the weighted sum of dark + bright. This analysis was performed at several other areas with similar results, before we successfully modelled the entire mosaic shown in Fig. 1.

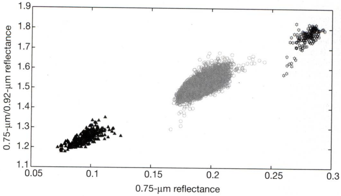


Figure 4 | Correlation of pyroxene absorption with material reflectance. The strength of the '1-µm' pyroxene absorption, measured by the ratio of reflectance at 0.75 µm to that at 0.92 µm, is strongly correlated with the reflectance in the continuum (0.75 µm), shown here for three example regions. Data are from the framing camera[2]. Greater ratio values correspond to stronger absorption bands. Dark (black filled circles) and bright (black open circles) areas cluster near the extremes of the plot, while background material (grey open circles) appears near the middle of this apparent mixing line. Dark material pixels are from Cornelia crater (9.3° S, 225.5° E), background material is from an area located between 5° S–17° N, 320° E–356° E and the bright material is from Tuccia crater (40° S, 197° E) in the Claudia coordinate system used by Dawn. If the entire surface of Vesta were treated here, rather than only example areas, the three clusters of points would merge into a continuum from least to greatest reflectance.

OH. Further, the Gamma Ray and Neutron Detector investigation reports[18] excess H (without identifying the molecular form) for certain regions that correspond with the broader darker areas shown in Fig. 1.

The most common group of differentiated meteorites, the howardite–eucrite–diogenites (HEDs), display reflectance spectra strongly suggesting that they originated from Vesta[8]. Three types of dark materials occur in the HEDs[19]: clasts of carbonaceous chondrite material in howardites, impact melts and shock-blackened materials in HED breccias, and fine-grained eucrites with quenched textures. These dark components occur in a few samples up to 60 per cent by volume but more typically compose a few per cent by volume of the rocks[20,21]. Although several types of dark material exist in the HEDs, only the carbonaceous material is consistent with the properties of the dark constituent of Vesta's surface.

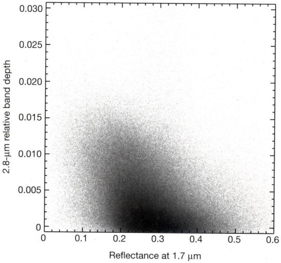


Figure 5 | Correlation of OH spectral absorption with material reflectance. Two-dimensional scatter plot from global observations of Vesta by VIR show a diffuse anti-correlation of the OH-related 2.8-µm-absorption band depth[9,15] and Vesta's reflectance at 1.7 µm (Fig. 1). The OH signature is correlated with the dark material in Vesta's surface.

It is plausible that enough dark material to match what we observe could be delivered on Vesta by low albedo (probably carbonaceous) asteroids over time. In addition to Vesta's small size and surface gravity, its location in the asteroid belt implies that most impacts occur at lower velocities, favouring the preservation of a larger fraction of the impacting material (see Supplementary Information). Analytical estimates we performed (see Supplementary Information) indicate that about 300 low-albedo asteroids with diameters between 1 km and 10 km could have impacted Vesta during the last 3.5 billion years, that is, once the asteroid belt assumed its present structure[22]. These impactors would deliver to the Vestan surface about (3–4) × 10^{18} g of low-albedo, probably carbonaceous material. This mass would be enough to wrap Vesta in a thick dark blanket 1–2 m thick. Impact gardening would occur, creating a dark crustal mixed zone to a depth of one to several kilometres. Further, it is well documented that the addition of small amounts of fine-grained opaque material to a semi-transparent material (such as Vesta's pyroxene-rich surface) reduces the reflectance of the mixture in a highly effective, nonlinear manner[23].

Thus, we see on Vesta a dramatic example of a common process that must affect many Solar System objects: that of contamination and alteration of indigenous surfaces owing to impact and retention and mixing of the impactor material. This process also provides a mechanism by which to transport hydrated and organic-rich materials to Vesta and other bodies. This hypothesis is consistent with evidence observed in the HED meteorites[19] and strengthens the connection between the HEDs and Vesta[8,19].

Received 16 June; accepted 29 August 2012.

1. Jaumann, R. *et al.* Vesta's shape and morphology. *Science* **336,** 687–690 (2012).
2. Reddy, V. *et al.* Color and albedo heterogeneity of Vesta from Dawn. *Science* **336,** 700–704 (2012).
3. Sandford, S. A. *et al.* Organics captured from comet 81P/Wild 2 by the Stardust Spacecraft. *Science* **314,** 1720–1724 (2006).
4. Sunshine, J. M. *et al.* Exposed water ice deposits on the surface of comet 9P/Temple 1. *Science* **311,** 1453–1455 (2006).
5. Gibb, E. L. *et al.* The organic composition of C/2001 A2 (LINEAR). II. Search for heterogeneity within a comet nucleus. *Icarus* **188,** 224–232 (2007).
6. Russell, C. T. *et al.* Dawn at Vesta: testing the protoplanetary paradigm. *Science* **336,** 684–686 (2012).
7. McCord, T. B., McFadden, L. A., Russell, C. T., Sotin, C. & Thomas, P. C. Ceres, Vesta, and Pallas: protoplanets, not asteroids. *Trans. AGU* **87,** 105–109 (2006).
8. McCord, T. B., Adams, J. B. & Johnson, T. V. Asteroid Vesta: spectral reflectivity and compositional implications. *Science* **168,** 1445–1447 (1970).
9. Degewij, J., Tedesco, E. F. & Zellner, B. H. Albedo and color contrasts on asteroid surfaces. *Icarus* **40,** 364–374 (1979).
10. Zellner, B. H. *et al.* Hubble Space Telescope images of asteroid 4 Vesta in 1994. *Icarus* **128,** 83–87 (1997).
11. Russell, C. T. & Raymond, C. A. The Dawn mission to Vesta and Ceres. *Space Sci. Rev.* **163,** 3–23 (2011).
12. De Sanctis, M. C. *et al.* Spectroscopic characterization of mineralogy and its diversity across Vesta. *Science* **336,** 697–700 (2012).
13. Combe, J.-Ph. *et al.* Analysis of OMEGA/Mars Express data hyperspectral data using a Multiple-Endmember Linear Spectral Unmixing Model (MELSUM): methodology and first results. *Planet. Space Sci.* **56,** 951–975 (2008).
14. Combe, J.-Ph. *et al.* Water and other volatiles on Vesta after the lunar case. *Lunar Planet. Sci. Conf.* **431,** 2463 (2012).
15. De Sanctis, M. C. *et al.* Detection of widespread hydrated materials on Vesta by VIR imaging spectrometer on board the Dawn mission. *Astrophys. J. Lett.* **758,** L36 (2012).
16. Pieters, C. M. *et al.* Character and spatial distribution of OH/H_2O on the surface of the Moon seen by M^3 on Chandrayaan-1. *Science* **326,** 568–572 (2009).
17. McCord, T. B. *et al.* Sources and physical processes responsible for OH/H_2O in the lunar soil as revealed by the Moon Mineralogy Mapper (M^3), *J. Geophys. Res.* **116,** E00G05 (2011).
18. Prettyman, T. H. *et al.* Elemental mapping by Dawn reveals exogenic H in Vesta's regolith. *Science* doi:10.1126/science.1225354.
19. McSween, H. Y., Mittlefehldt, D. W., Beck, A. W., Mayne, R. G. & McCoy, T. J. HED meteorites and their relationship to the geology of Vesta and the Dawn mission. *Space Sci. Rev.* **163,** 141–174 (2011).
20. Zolensky, M. E., Weisberg, M. K., Buchanan, P. C. & Mittlefehldt, D. W. Mineralogy of carbonaceous chondrite clasts in HED achondrites and the Moon. *Meteorit. Planet. Sci.* **31,** 518–537 (1996).
21. Beck, A. W., Welten, K. C., McSween, H. Y., Viviano, C. E. & Caffee, M. W. Petrologic and textural diversity among the PCA 02 howardite group, one of the largest pieces of the Vestan surface. *Meteorit. Planet. Sci.* **47,** 947–969 (2012).

22. Coradini, A., Turrini, D., Federico, C. & Magni, G. Vesta and Ceres: crossing the history of the Solar System. *Space Sci. Rev.* **163,** 25–40 (2011).
23. Clark, R. N. Spectral properties of mixtures of montmorillonite and dark grains—implications for remote sensing minerals containing chemically and physically adsorbed water. *J. Geophys. Res.* **88,** 10635–10641 (1983).

Supplementary Information is available in the online version of the paper.

Acknowledgements This research was supported by the NASA Dawn Project under contract from UCLA, by the NASA Dawn at Vesta Participating Scientist program, the Italian Space Agency, the Max Planck Institute for Solar System Research, and the Germany Aerospace Agency (DLR). We acknowledge the support of the Dawn Science, Instrument and Operations Teams.

Author Contributions T.B.McC. designed the study, directed the research and wrote the manuscript; J.-P.C., J-Y.L. and H.Y.McS. were the major contributors helping to design the study, direct the research and write the manuscript. D.T., T.B.McC. and J.-P.C. wrote the Supplementary Information. J.-P.C. performed the spectral unmixing of the VIR mapping spectrometer data and the correlation of OH with albedo. J.-P.C. and F.T. produced global maps from the VIR mapping spectrometer data. V.R. performed the correlation of albedo with band strength. L.L.C. and R.J. produced global mapping of Vesta with Framing Camera images. H.Y.McS., V.R. and D.W.M. contributed to the laboratory analysis of Vestan meteorites. J.-P.C., E.P., E.A., M.C.D.S. and A.L. contributed to the location of dark materials and to the spectral analysis of hydroxyl-rich areas from the VIR mapping spectrometer data. R.J. provided the location of dark materials from Framing Camera images. D.T. provided estimates of amounts of dark material coming from carbonaceous chondrite meteorites. J.-P.C., J-Y.L., F.C., E.P. and A.L. contributed to the photometric correction for the VIR mapping spectrometer. J-Y.L., S.E.S. and L.L.C. contributed to the photometric correction of Framing Camera images. F.T. and J.-P.C. investigated the need for thermal emission correction of the VIR mapping spectrometer. D.A.W. contributed much of the geologic context of dark material, developed example site descriptions, and wrote or edited geological context materials. D.T.B. participated in developing the dark-material interpretation and edited the manuscript. C.M.P. participated in providing context for the mixing effects on the resulting surface material and integrating space-weathering concepts. M.C.D.S. is the Principal Investigator of the VIR mapping spectrometer and helped to provide the VIR data. A.N. is the Principal Investigator of the Framing Camera and helped to provide the Framing Camera data. M.T.C. and F.C. analysed the thermal properties of the surface of Vesta. D.W.M. located bright materials from Framing Camera images. E.A., F.C., F.T., A.L., M.C.D.S. and J.-P.C. contributed to the calibration of the VIR mapping spectrometer. A.N., S.E.S. and V.R. contributed to the calibration of the Framing Camera. E.A. contributed to the sequencing of data acquisition for the VIR mapping spectrometer. S.E.S. contributed to the sequencing of data acquisition for the Framing Camera. H.H. contributed to the design and development of the Framing Camera. A.W.B., U.C., H.U.K., B.W.D. & J.M.S. participated in the working group studying the bright material. C.A.R. is the deputy Principal Investigator of the Dawn mission. C.T.R. is the Principal Investigator of the Dawn mission. All authors discussed the results and commented on the manuscript.

Author Information Data from the Dawn mission are publicly available through the NASA Planetary Data System. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.B.M. (tmccord@bearfightinstitute.com) or J.-P.C. (jean-philippe_combe@bearfightinstitutute.com).

LETTER

doi:10.1038/nature11596

Observation of spatially ordered structures in a two-dimensional Rydberg gas

Peter Schauß[1], Marc Cheneau[1], Manuel Endres[1], Takeshi Fukuhara[1], Sebastian Hild[1], Ahmed Omran[1], Thomas Pohl[2], Christian Gross[1], Stefan Kuhr[1,3] & Immanuel Bloch[1,4]

The ability to control and tune interactions in ultracold atomic gases has paved the way for the realization of new phases of matter. So far, experiments have achieved a high degree of control over short-range interactions, but the realization of long-range interactions has become a central focus of research because it would open up a new realm of many-body physics. Rydberg atoms are highly suited to this goal because the van der Waals forces between them are many orders of magnitude larger than those between ground-state atoms[1]. Consequently, mere laser excitation of ultracold gases can cause strongly correlated many-body states to emerge directly when atoms are transferred to Rydberg states. A key example is a quantum crystal composed of coherent superpositions of different, spatially ordered configurations of collective excitations[2–5]. Here we use high-resolution, *in situ* Rydberg atom imaging to measure directly strong correlations in a laser-excited, two-dimensional atomic Mott insulator[6]. The observations reveal the emergence of spatially ordered excitation patterns with random orientation, but well-defined geometry, in the high-density components of the prepared many-body state. Together with a time-resolved analysis, this supports the description of the system in terms of a correlated quantum state of collective excitations delocalized throughout the gas. Our experiment demonstrates the potential of Rydberg gases to realize exotic phases of matter, thereby laying the basis for quantum simulations of quantum magnets with long-range interactions.

The strongly enhanced interaction between Rydberg atoms makes them unique building blocks for a variety of applications, ranging from quantum optics and quantum information processing[1,7,8] to engineering of exotic quantum many-body phases[9–11]. For the last purpose, two main ideas have been explored theoretically. First, the weak admixing of a Rydberg state to the atomic ground state using off-resonant laser coupling has been suggested as a way to benefit from the long-range interactions without persistent population of the Rydberg state[9,10,12]. Second, direct laser excitation leads to the formation of a gas of Rydberg excitations, also called a Rydberg gas. This strongly correlated system[13] can exhibit highly non-classical states characterized by the coherent superposition of ordered structures in the spatial distribution of the Rydberg excitations[2–5,14–16]. Here the excitation dynamics proceeds on a timescale of a few microseconds, on which the atoms can be considered frozen in space, representing strongly interacting effective spins. At the heart of the formation of such correlated states lies the dipole blockade effect[1,7,8] that prevents simultaneous Rydberg excitation of two close-by atoms[17–21]. Recent experiments using two trapped atoms have shown how this blockade effect can be used to implement fast two-qubit quantum gates[22,23]. In larger ultracold atomic ensembles, the coherence of the collective excitation has been demonstrated[24–26] and evidence for strong correlations could be found by observing universal scaling laws for the number of excited Rydberg atoms[27,28]. However, direct measurements of spatial ordering have remained an outstanding challenge. Important steps in this direction were recently explored using a field-ion-microscope[20], allowing the measurement of the blockade radius in a three-dimensional Rydberg gas. Recent theoretical work, on the other hand, has proposed detection schemes with potential resolution below the blockade radius, based on conditional Raman transfer[29] or electromagnetically induced transparency[30].

Here we demonstrate an alternative approach that permits direct imaging of Rydberg excitation patterns, and precise measurements of correlation functions. This allows us to probe the underlying constituents of the excited many-body state, revealing the spatial ordering of the high-density components. Two key advances form the basis of our observations. First, a two-dimensional atomic Mott insulator provides a dense and well-ordered initial system that maximizes coherence times during the excitation dynamics. Second, we developed an all-optical technique to image individual Rydberg atoms *in situ* with high spatial and temporal resolution.

The physical system considered here is a two-dimensional gas of rubidium atoms trapped in a rotationally invariant harmonic confinement potential and pinned in a square optical lattice. The gas was prepared deep in the Mott-insulating phase, ensuring uniform filling with one atom per site within a disk of radius $R \approx \sqrt{N_{\mathrm{at}} a_{\mathrm{lat}}^2/\pi}$, where N_{at} is the total number of atoms and a_{lat} the lattice spacing. The atoms were initially in their electronic ground state, $|g\rangle$, and then resonantly coupled to a Rydberg state, $|e\rangle$. In the interaction picture, the internal dynamics of the atoms is governed by the many-body Hamiltonian:

$$\hat{H} = \frac{\hbar\Omega}{2}\sum_{i}\left(\hat{\sigma}_{eg}^{(i)} + \hat{\sigma}_{ge}^{(i)}\right) + \sum_{i \neq j}\frac{V_{ij}}{2}\hat{\sigma}_{ee}^{(i)}\hat{\sigma}_{ee}^{(j)} \qquad (1)$$

Here, the vectors $i = (i_x, i_y)$ label the lattice sites in the plane. The first term in this Hamiltonian describes the coherent coupling of the ground and excited states with Rabi frequency Ω, where $\hat{\sigma}_{ge}^{(i)} = |e_i\rangle\langle g_i|$ and $\hat{\sigma}_{eg}^{(i)} = |g_i\rangle\langle e_i|$ are the local transition operators. The second term is the van der Waals interaction potential between two atoms in the Rydberg state. In our case it is repulsive and takes the asymptotic form $V_{ij} = -C_6 / r_{ij}^6$, with the van der Waals coefficient $C_6 < 0$ and $r_{ij} = a_{\mathrm{lat}}|i - j|$ the distance between the two atoms at sites i and j. The projection operator $\hat{\sigma}_{ee}^{(i)} = |e_i\rangle\langle e_i|$ measures the population of the Rydberg state at site i. This model is valid as long as the mechanical motion of the atoms and all decoherence effects can be neglected (Supplementary Information).

The dynamics of this strongly correlated system can be understood intuitively from its energy spectrum in the absence of optical driving. It is instructive to group the large number of many-body states, $2^{N_{\mathrm{at}}}$, according to the number of Rydberg excitations, N_{e}, contained in each state (Fig. 1a). All singly excited states ($N_{\mathrm{e}} = 1$) with different positions of the Rydberg atom have identical energies and form a N_{at}-fold degenerate manifold. For multiply excited states ($N_{\mathrm{e}} > 1$), this degeneracy is lifted by the strong van der Waals interaction, giving rise to a broad energy band (Fig. 1a). Starting from the ground state, the creation of

---

[1]Max-Planck-Institut für Quantenoptik, 85748 Garching, Germany. [2]Max-Planck-Institut für Physik komplexer Systeme, 01187 Dresden, Germany. [3]University of Strathclyde, Department of Physics, SUPA, Glasgow G4 0NG, UK. [4]Ludwig-Maximilians-Universität, Fakultät für Physik, 80799 München, Germany.

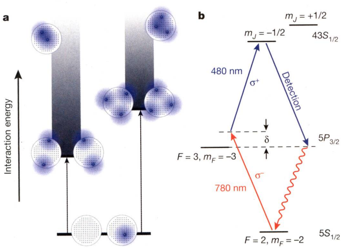


Figure 1 | Schematics of the many-body excitation. a, Energy spectrum in the absence of optical driving. States with more than one excitation form a broad energy band (shown as a grey shading for $N_e = 2$ on the left and $N_e = 3$ on the right) above the degenerate manifold comprising the ground state and all singly excited states. For each excitation number $N_e > 1$, the states with lowest energy correspond to spatially ordered configurations, which maximize the distance between the Rydberg excitations. The minimal interaction energy (black arrows) is determined by the finite system size and increases with N_e. Possible spatial configurations of the excitations (blue dots) in the initial Mott-insulating state (black dots) are shown schematically as circular insets next to their respective interaction energy. The blockade radius is depicted by the blue shaded disk around the excitation. **b,** Simplified level scheme of ^{87}Rb, showing the transitions used for the Rydberg excitation and detection. See text for details.

the first excitation is resonant, while the sequential coupling to many-body states with larger number of excitations is rapidly detuned by the interactions. In fact, the rapid variation of the van der Waals potential with distance prevents the excitation of all those states where Rydberg atoms are separated by less than the blockade radius, R_b, defined by $\hbar\Omega = -C_6/R_b^6$. The existence of this exclusion radius is expected to have a striking consequence: whereas the total many-body state exhibits finite-range correlations on a scale of R_b (ref. 13), its high-density components with a Rydberg density close to $1/R_b^2$ should display a crystalline structure, meaning that the position of the Rydberg atoms is correlated over a distance comparable to the system size.

The excitation dynamics of all configurations should occur in an entirely coherent fashion, resulting in highly non-classical many-body states. The approximate rotational symmetry of our system leads to symmetric superpositions of all microscopic configurations with different orientation but identical relative positions of the Rydberg atoms. Also, as the coupling addresses all states within an energy range $\sim\hbar\Omega$, it produces a coherent superposition of many-body states with different numbers of excitations and slightly different separation between the Rydberg atoms (Fig. 1a). This collective nature of the excited many-body states dramatically changes the timescale on which their dynamics occurs. The coupling strength to the state with a single excitation is enhanced by a factor $\sqrt{N_{at}} \gg 1$ (ref. 8) and the coupling to states with $N_e > 1$ is similarly enhanced, with N_{at} replaced by the number of energetically accessible configurations in each N_e-manifold[3].

Our experiments began with the preparation of a two-dimensional degenerate gas of 150–390 ^{87}Rb atoms confined to a single antinode of a vertical (z-axis) optical lattice[31]. The gas was brought deep into the Mott-insulating phase by adiabatically turning on a square optical lattice with period $a_{lat} = 532$ nm in the x–y plane. Within the system radius, $R = 3.5$ μm to 5 μm, the probability of a lattice site being occupied by a single atom was typically 80%. The atoms were then initialized in the hyperfine ground state $|g\rangle \equiv |5S_{1/2}, F = 2, m_F = -2\rangle$ and coupled to the Rydberg state $|e\rangle \equiv |43S_{1/2}, m_J = -1/2\rangle$, using the standard notation for the fine and hyperfine structure (Fig. 1b). The coupling was achieved through a two-photon process via the intermediate state $|5P_{3/2}, F = 3, m_F = -3\rangle$ using lasers of wavelengths 780 nm and 480 nm and σ^- and σ^+ polarization, respectively (Fig. 1b and Methods). The resulting two-photon Rabi frequency was $\Omega/(2\pi) = 170(20)$ kHz (the number in parentheses denotes the uncertainty of the last digit), yielding a blockade radius of $R_b = 4.9(1)$ μm. Following the initial preparation, we suddenly switched on the excitation lasers and let the system evolve for a variable duration t. After the excitation pulse, we detected the Rydberg excitations by first removing all atoms in the ground state with a resonant laser pulse, then de-exciting the Rydberg atoms to the ground state via stimulated emission towards the intermediate state (Fig. 1b and Methods) and finally recording their position using high-resolution fluorescence imaging[31]. The accuracy of the measurement was limited by the probability, 75(10)%, of detecting a Rydberg atom and by a background signal due to on average 0.2(1) non-removed ground state atoms per picture (Supplementary Information). The spatial resolution of our detection technique is limited to about one lattice site by the residual motion of the atoms in the Rydberg state before de-excitation (Supplementary Information). Repeating the experiment many times allowed for sampling the different spatial configurations of Rydberg atoms constituting the many-body state and to measure their respective statistical weight.

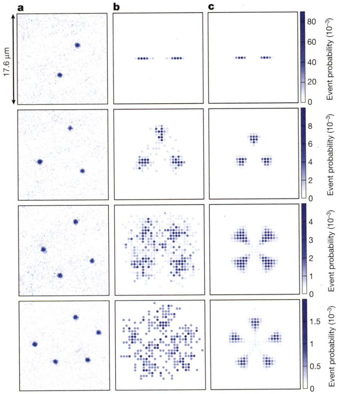


Figure 2 | Spatially ordered components of the many-body states. Spatial distribution of excitations for the observed microscopic configurations sorted according to their number of excitations, $N_e = 2$–5 (top to bottom). **a,** Examples of false-colour fluorescence images in which de-excited Rydberg atoms are directly visible as dark-blue spots. **b,** Histograms of the spatial distribution of Rydberg atoms obtained after centring and aligning the individual microscopic configurations to a reference axis (Methods). The initial atom distribution had a diameter of 7.2(8) μm and 10.8(8) μm for $N_e = 2$–3 and $N_e = 4$–5, respectively. **c,** Theoretical prediction from numerical simulations of the excitation dynamics governed by the many-body Hamiltonian of equation (1) for the same conditions as in the experiment (Supplementary Information). Colour scale at right of each row applies only to the sub-panels of **b** and **c** in that row.

In Fig. 2a we show typical images of microscopic configurations with $N_e = 2$–5. In order to analyse the structure of the many-body state, we group the individual images according to their number of excitations and determine the spatial distributions of the excitations, $\rho_e(i) = \left\langle \hat{\sigma}_{ee}^{(i)} \right\rangle$, where $\langle\cdot\rangle$ denotes the average from repeated measurements. These distributions display a typical ring-shaped profile (Supplementary Fig. 1), which results from the blockade effect and from the rotational symmetry of the system. Spatially ordered structures become visible once each microscopic configuration has been centred and aligned to a fixed reference axis (Fig. 2b and Methods).

For our smallest sample ($R \approx 3.5\,\mu$m), we observe strong correlations between $N_e = 2$ excitations that are separated by a distance of $\sim 6\,\mu$m, due to the interaction blockade. In the same data set, configurations with $N_e = 3$ show an arrangement on an equilateral triangle, revealing both strong radial and azimuthal ordering. These correlations persist for larger numbers of Rydberg excitations, which we can prepare in larger samples ($R \approx 5\,\mu$m). They form quadratic and pentagonal configurations for $N_e = 4$ and $N_e = 5$, respectively. However, since their interaction energy is larger, these states are populated only with low probability, leading to a reduced signal-to-noise ratio. Our experimental data are in good agreement with numerical simulations of the many-body dynamics according to the Hamiltonian of equation (1), for the same atom numbers, temperature and laser parameters as in the experiment (Fig. 2c and Supplementary Information). These simulations are based on a truncation of the underlying Hilbert space, exploiting the dipole blockade, and neglect any dissipative effects (ref. 3 and Supplementary Information). The spatial distributions of excitations provided by the simulation reproduce all the features observed in the experiment. The only apparent discrepancy is the overall slightly larger size of the measured structures, which can be attributed to the spatial resolution of our detection method, as discussed below.

For a more quantitative analysis of spatial correlations, we also measured the pair correlation function (Fig. 3a)

$$g^{(2)}(r) = \frac{\sum_{i \neq j} \delta_{r,r_{ij}} \left\langle \hat{\sigma}_{ee}^{(i)} \hat{\sigma}_{ee}^{(j)} \right\rangle}{\sum_{i \neq j} \delta_{r,r_{ij}} \left\langle \hat{\sigma}_{ee}^{(i)} \right\rangle \left\langle \hat{\sigma}_{ee}^{(j)} \right\rangle} \qquad (2)$$

which characterizes the occurrence of two excitations separated by a distance r. Here $\delta_{r,r_{ij}}$ is the Kronecker symbol that restricts the sum to sites (i, j) for which $r_{ij} = r$. In contrast to the spatial distributions presented above, the average is now taken over all values of N_e. The pair correlation function $g^{(2)}(r)$ shows a strong suppression at distances smaller than $r = 4.8(2)\,\mu$m, which coincides with the expected blockade radius $R_b = 4.9(1)\,\mu$m. Moreover, we find a clear peak at $r = 5.6(2)\,\mu$m and evidence for weak oscillations extending to the boundaries of our system. This indicates that the overall many-body state only exhibits finite-range correlations. Our theoretical calculation of $g^{(2)}(r)$ (grey line in Fig. 3a) exhibits similar features, but shows more pronounced oscillations and vanishes perfectly within the blockade radius. These discrepancies can be attributed to several imperfections of the detection technique. The sharp peak at short distances $r \lesssim 1\,\mu$m results from hopping of single atoms to adjacent sites during fluorescence imaging with a small probability of approximately 1%, which is falsely detected as two neighbouring excitations. The non-zero value of $g^{(2)}(r)$ for distances $r \lesssim 3\,\mu$m arises from the imperfect removal of the ground-state atoms. Finally, the shift and slight broadening of the peak in the correlation function is attributed to the residual motion of the Rydberg atoms before imaging (Supplementary Information). When we take account of these independently characterized effects in the theoretical calculations (green line in Fig. 3a), we recover excellent agreement with the measurements.

Because our system size is comparable to the blockade radius, the excitations in states with $N_e > 1$ are localized along the circumference of the system. We characterize the resulting angular order by introducing an azimuthal correlation function that reflects the probability of finding two excitations with a relative angle $\Delta\phi$ measured with respect to the centre of mass of the distribution of excitations:

$$\tilde{g}^{(2)}(\Delta\phi) = \int \frac{\mathrm{d}\phi}{2\pi} \frac{\langle \hat{n}(\phi)\hat{n}(\phi+\Delta\phi) \rangle}{\langle \hat{n}(\phi) \rangle \langle \hat{n}(\phi+\Delta\phi) \rangle} \qquad (3)$$

Here $\hat{n}(\phi) = \sum_i \delta_{\phi,\phi_i} \hat{\sigma}_{ee}^{(i)}$ is the azimuthal distribution of excitations, with (r_i, ϕ_i) the polar coordinates of site i. As can be seen in Fig. 3b, the spatially ordered structure is clearly visible as correlations at relative angles $\Delta\phi = \nu \times 360°/N_e$, with $\nu = 1, 2, \ldots, N_e$, even for the largest excitation numbers.

We finally analyse the many-body excitation dynamics of the system. In Fig. 4a we show the time evolution of the average number of Rydberg excitations, $\bar{N}_e = \sum_i \left\langle \hat{\sigma}_{ee}^{(i)} \right\rangle$, which quickly saturates to a

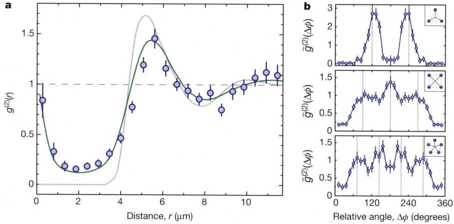


Figure 3 | Correlation functions of Rydberg excitations. a, Pair correlation function. The blockade effect results in a strong suppression of the probability of finding two excitations separated by a distance less than the blockade radius $R_b = 4.9(1)\,\mu$m. Moreover, we observe a peak at $r \approx 5.6\,\mu$m and a weak oscillation at larger distances. The initial atom distribution had a diameter of $10.8(8)\,\mu$m. The experimental data (blue circles) are compared to the theoretical prediction both taking into account the independently characterized imperfections of our detection method (green line) and disregarding these imperfections (grey line). The dashed line marks the value of $g^{(2)}$ in the absence of correlations. Error bars, s.e.m. of $g^{(2)}(r)$. **b**, Azimuthal correlation function. The spatially ordered structure of the high-density components is best visible in the angular correlations around the centre of mass of the distribution of excitations, characterized by the correlation function $\tilde{g}^{(2)}(\Delta\phi)$ defined in equation (3). By construction, this function is symmetric around 180°. Correlations are observed at the angles expected for the respective configurations shown in the insets. The peaks close to 180° are more pronounced because the centre of mass of a configuration is likely to lie close to the intersections of the diagonals, owing to the blockade effect. Error bars, s.e.m.

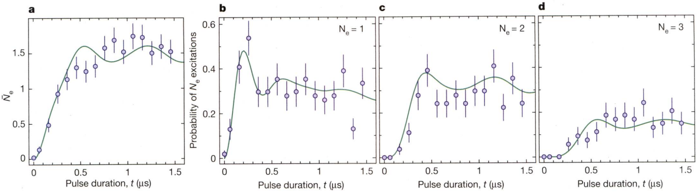


Figure 4 | Time evolution of the number of Rydberg excitations. a, Average number of detected Rydberg atoms as a function of the excitation pulse duration. Error bars, s.e.m. **b–d,** Time evolution of the probability of observing $N_e = 1$ (**b**), $N_e = 2$ (**c**) and $N_e = 3$ (**d**) Rydberg excitations. The experimental data (blue circles) are compared to the theoretical prediction (green line), which is based on initial ground-state atom distributions observed in the experiment and neglects all decoherence effects. It takes into account the finite detection efficiency as a free parameter (75%). Error bars, s.e.m.

small value $\bar{N}_e \approx 1.5$, much smaller than the total number of atoms in the system, $N_{at} = 150(30)$. The saturation is reached in ~500 ns, a factor of ten faster than the Rabi period $2\pi/\Omega$, owing to the collective enhancement of the optical coupling strength. The probability of observing N_e Rydberg excitations shows a similar saturation profile for each excitation number N_e (Fig. 4b–d), but on a timescale that increases with N_e, from about 200 ns for $N_e = 1$ to about 600 ns for $N_e = 3$. This can be attributed to the variation of the collective enhancement factor associated with the number of energetically accessible microscopic configurations for a given N_e. The theoretical excitation dynamics corresponding to the Hamiltonian, equation (1), shows remarkable agreement with the experimental data when the finite detection efficiency is included. This provides evidence that the dynamics observed in the experiment is coherent, as expected on these timescales, which are much shorter than the lifetime of the Rydberg state of 25(5) μs in the lattice and the timescale of other decoherence effects (Supplementary Information). The absence of high-contrast Rabi oscillations in the time evolution of the average number of Rydberg excitations is caused by the strong dephasing between many-body states with different interaction energies arising from the different spatial distribution of excitations. However, remnant signatures of Rabi oscillations can still be observed. In particular, the population of the singly excited states shows a peak around $t = 200(50)$ ns (Fig. 4b), which matches the π-pulse time of the enhanced Rabi frequency $\pi/(\sqrt{N_{at}}\Omega) = 240(40)$ ns. Further evidence for the coherence of the dynamics can be found in the spatially resolved analysis of the excitation dynamics (Supplementary Information).

We have characterized the strongly correlated excitation dynamics of a resonantly driven Rydberg gas using optical detection with unprecedented spatial resolution, and observed ordered Rydberg excitation patterns in the high-density components of the many-body states produced. One future challenge lies in the deterministic preparation of ground-state Rydberg crystals with a well-defined number of excitations via adiabatic sweeps of the laser parameters[3,4,14]. Together with the demonstrated imaging technique, this would enable precise studies of quantum phase transitions in long-range interacting quantum systems on the microscopic level[2–4,15]. Also, by combining the dipole blockade effect with the single-atom addressing already demonstrated in our experimental set-up, it should be possible to engineer mesoscopic quantum gates[32]; such a system could serve as an experimental 'toolbox' for digital quantum simulations of a broad class of spin models, including such fundamental systems as Kitaev's toric code[33].

METHODS SUMMARY

Rydberg excitation and detection scheme. The two excitation laser beams were counterpropagating along the z axis, with an intermediate-state detuning $\delta/(2\pi) = 742(2)$ MHz (Fig. 1b). During the sequence, a magnetic offset field of $B \approx 30$ G along the z axis defined the quantization axis. The excitation pulse was performed by switching the laser at 780 nm while the laser at 480 nm was on. The temporal resolution of our measurement was thus set by the rise time of the 780 nm light, which was ~40 ns. Immediately after the excitation pulse, we used near-resonant circularly polarized laser beams to drive the transitions $|5S_{1/2}, F=1\rangle \rightarrow |5P_{3/2}, F=2\rangle$ and $|5S_{1/2}, F=2\rangle \rightarrow |5P_{3/2}, F=3\rangle$ and remove all ground-state atoms, with a fidelity of 99.9% in 10 μs. Subsequently, the Rydberg atoms were stimulated down to the ground state by resonantly driving the $|43S_{1/2}, m_J=-1/2\rangle \rightarrow |5P_{3/2}, F=3, m_F=-3\rangle$ transition for 2 μs.

Computation of the histograms. The histograms shown in Fig. 2b are based on the digitized atom distribution reconstructed from the raw images[31]. They reflect the Rydberg atom distribution in a region of interest covering a disk of radius $R_{max} = 1.5 \times R$. Each individual image was aligned in the following way. First, we set the origin of the coordinate system to the centre of mass of the atom distribution. Then, for each atom we determined the angle between its position vector and a reference axis, and rotated the image about the origin by the mean value of these angles (repeating this operation would leave the configuration unchanged). The histograms contain data taken at different evolution times up to 4 μs, as we found no significant temporal dependence of the excitation patterns. The theoretical calculations used the same parameters as in the experiment (including temperature and atom number distribution of the initial state) and followed the same procedure to determine the Rydberg atom densities. Both the experimental and theoretical histograms were normalized such that the value at each bin represents the probability of observing a microscopic configuration with a Rydberg atom located at this position.

Received 11 July; accepted 13 September 2012.

1. Saffman, M., Walker, T. & Mølmer, K. Quantum information with Rydberg atoms. *Rev. Mod. Phys.* **82,** 2313–2363 (2010).
2. Weimer, H., Löw, R., Pfau, T. & Büchler, H. P. Quantum critical behavior in strongly interacting Rydberg gases. *Phys. Rev. Lett.* **101,** 250601 (2008).
3. Pohl, T., Demler, E. & Lukin, M. D. Dynamical crystallization in the dipole blockade of ultracold atoms. *Phys. Rev. Lett.* **104,** 043002 (2010).
4. Schachenmayer, J., Lesanovsky, I., Micheli, A. & Daley, A. J. Dynamical crystal creation with polar molecules or Rydberg atoms in optical lattices. *N. J. Phys.* **12,** 103044 (2010).
5. Gärttner, M., Heeg, K. P., Gasenzer, T. & Evers, J. Dynamical formation of floating Rydberg excitation crystals. Preprint at http://arXiv.org/abs/1203.2884v2 (2012).
6. Bloch, I., Dalibard, J. & Zwerger, W. Many-body physics with ultracold gases. *Rev. Mod. Phys.* **80,** 885–964 (2008).
7. Jaksch, D., Cirac, J. I., Zoller, P., Côté, R. & Lukin, M. D. Fast quantum gates for neutral atoms. *Phys. Rev. Lett.* **85,** 2208–2211 (2000).
8. Lukin, M. *et al.* Dipole blockade and quantum information processing in mesoscopic atomic ensembles. *Phys. Rev. Lett.* **87,** 037901 (2001).
9. Pupillo, G., Micheli, A., Boninsegni, M., Lesanovsky, I. & Zoller, P. Strongly correlated gases of Rydberg-dressed atoms: quantum and classical dynamics. *Phys. Rev. Lett.* **104,** 223002 (2010).
10. Henkel, N., Nath, R. & Pohl, T. Three-dimensional roton excitations and supersolid formation in Rydberg-excited Bose-Einstein condensates. *Phys. Rev. Lett.* **104,** 195302 (2010).
11. Cinti, F. *et al.* Supersolid droplet crystal in a dipole-blockaded gas. *Phys. Rev. Lett.* **105,** 135301 (2010).
12. Honer, J., Weimer, H., Pfau, T. & Büchler, H. Collective many-body interaction in Rydberg dressed atoms. *Phys. Rev. Lett.* **105,** 160404 (2010).
13. Robicheaux, F. & Hernández, J. Many-body wave function in a dipole blockade configuration. *Phys. Rev. A* **72,** 063403 (2005).

14. van Bijnen, R. M. W., Smit, S., van Leeuwen, K. H., Vredenbregt, E. J. D. & Kokkelmans, S. J. J. M. F. Adiabatic formation of Rydberg crystals with chirped laser pulses. *J. Phys. B* **44,** 184008 (2011).
15. Lesanovsky, I. Many-body spin interactions and the ground state of a dense Rydberg lattice gas. *Phys. Rev. Lett.* **106,** 025301 (2011).
16. Höning, M., Muth, D., Petrosyan, D. & Fleischhauer, M. Steady-state crystallization of Rydberg excitations in an optically driven lattice gas. Preprint at http://arXiv.org/abs/1208.2911 (2012).
17. Tong, D. *et al.* Local blockade of Rydberg excitation in an ultracold gas. *Phys. Rev. Lett.* **93,** 063001 (2004).
18. Urban, E. *et al.* Observation of Rydberg blockade between two atoms. *Nature Phys.* **5,** 110–114 (2009).
19. Gaëtan, A. *et al.* Observation of collective excitation of two individual atoms in the Rydberg blockade regime. *Nature Phys.* **5,** 115–118 (2009).
20. Schwarzkopf, A., Sapiro, R. & Raithel, G. Imaging spatial correlations of Rydberg excitations in cold atom clouds. *Phys. Rev. Lett.* **107,** 103001 (2011).
21. Dudin, Y. O. & Kuzmich, A. Strongly interacting Rydberg excitations of a cold atomic gas. *Science* **336,** 887–889 (2012).
22. Wilk, T. *et al.* Entanglement of two individual neutral atoms using Rydberg blockade. *Phys. Rev. Lett.* **104,** 010502 (2010).
23. Isenhower, L. *et al.* Demonstration of a neutral atom controlled-NOT quantum gate. *Phys. Rev. Lett.* **104,** 010503 (2010).
24. Raitzsch, U. *et al.* Echo experiments in a strongly interacting Rydberg gas. *Phys. Rev. Lett.* **100,** 013002 (2008).
25. Reetz-Lamour, M., Amthor, T., Deiglmayr, J. & Weidemüller, M. Rabi oscillations and excitation trapping in the coherent excitation of a mesoscopic frozen Rydberg gas. *Phys. Rev. Lett.* **100,** 253001 (2008).
26. Dudin, Y. O., Li, L., Bariani, F. & Kuzmich, A. Observation of coherent many-body Rabi oscillations. *Nature Phys.* http://dx.doi.org/10.1038/nphys2413 (published online, 9 September 2012).
27. Löw, R. *et al.* Universal scaling in a strongly interacting Rydberg gas. *Phys. Rev. A* **80,** 033422 (2009).
28. Viteau, M. *et al.* Rydberg excitations in Bose-Einstein condensates in quasi-one-dimensional potentials and optical lattices. *Phys. Rev. Lett.* **107,** 060402 (2011).
29. Olmos, B., Li, W., Hofferberth, S. & Lesanovsky, I. Amplifying single impurities immersed in a gas of ultracold atoms. *Phys. Rev. A* **84,** 041607(R) (2011).
30. Günter, G. *et al.* Interaction enhanced imaging of individual Rydberg atoms in dense gases. *Phys. Rev. Lett.* **108,** 013002 (2012).
31. Endres, M. *et al.* Observation of correlated particle-hole pairs and string order in low-dimensional Mott insulators. *Science* **334,** 200–203 (2011).
32. Müller, M., Lesanovsky, I., Weimer, H., Büchler, H. P. & Zoller, P. Mesoscopic Rydberg gate based on electromagnetically induced transparency. *Phys. Rev. Lett.* **102,** 170502 (2009).
33. Weimer, H., Müller, M., Lesanovsky, I., Zoller, P. & Büchler, H. P. A Rydberg quantum simulator. *Nature Phys.* **6,** 382–388 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R. Löw for discussions. We acknowledge funding by MPG, DFG, EU (NAMEQUAM, AQUTE, Marie Curie Fellowship to M.C.) and JSPS (Postdoctoral Fellowship for Research Abroad to T.F.).

Author Contributions All authors contributed extensively to the work presented in this paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.S. (peter.schauss@mpq.mpg.de).

doi:10.1038/nature11513

Fluvial response to abrupt global warming at the Palaeocene/Eocene boundary

Brady Z. Foreman[1], Paul L. Heller[1] & Mark T. Clementz[1]

Climate strongly affects the production of sediment from mountain catchments as well as its transport and deposition within adjacent sedimentary basins[1–3]. However, identifying climatic influences on basin stratigraphy is complicated by nonlinearities, feedback loops, lag times, buffering and convergence among processes within the sediment routeing system[3,4]. The Palaeocene/Eocene thermal maximum (PETM) arguably represents the most abrupt and dramatic instance of global warming in the Cenozoic era and has been proposed to be a geologic analogue for anthropogenic climate change[5]. Here we evaluate the fluvial response in western Colorado to the PETM. Concomitant with the carbon isotope excursion marking the PETM we document a basin-wide shift to thick, multistoried, sheets of sandstone characterized by variable channel dimensions, dominance of upper flow regime sedimentary structures, and prevalent crevasse splay deposits. This progradation of coarse-grained lithofacies matches model predictions for rapid increases in sediment flux and discharge[1,3], instigated by regional vegetation overturn[5,6] and enhanced monsoon precipitation[7,8]. Yet the change in fluvial deposition persisted long after the approximately 200,000-year-long PETM[9] with its increased carbon dioxide levels in the atmosphere, emphasizing the strong role the protracted transmission of catchment responses to distant depositional systems has in constructing large-scale basin stratigraphy. Our results, combined with evidence for increased dissolved loads[10] and terrestrial clay export[5,11,12] to world oceans, indicate that the transient hyper-greenhouse climate of the PETM may represent a major geomorphic 'system-clearing event'[13], involving a global mobilization of dissolved and solid sediment loads on Earth's surface.

During the PETM, an extreme global warming event that occurred about 56 million years ago[5], mean annual temperatures increased by 5°–8° C, precipitation and vegetation patterns dramatically altered worldwide, and both atmospheric and oceanic circulation was perturbed[5,6,8]. The warming was associated with a massive exogenic pulse of isotopically light carbon into Earth's oceans and atmosphere, recorded as a major negative carbon isotope excursion in a suite of organic and carbonate substrates hosted in marine and terrestrial strata[5]. Although more than 4,000 petagrams of carbon (Pg C) were released in less than 10,000 years (10 kyr), a high atmospheric partial pressure of CO$_2$ ($p_{\mathrm{CO_2}}$ exceeding 1,200 p.p.m.; ref. 5) persisted for an additional 190 kyr or so before being sequestered[5,9].

Here we characterize shifts in the nature of fluvial deposition spanning the PETM within the intermontane Piceance Creek basin of western Colorado, USA (Fig. 1), which formed during the Laramide orogeny[14]. Palaeocene and early Eocene deposition[15–17] is represented by the Wasatch formation, and is separated from base to top into the Atwell Gulch, Molina and Shire members[16]. Our new δ^{13}C record using dispersed organic carbon documents an approximately 3.0‰ excursion from background values of about −23.0‰ (Vienna Pee-Dee Belemnite standard, VPDB) (Fig. 2; Supplementary Data), constrained between pollen and mammalian fossil localities characteristic of the latest Palaeocene and earliest Eocene epochs[15–17] (Supplementary Discussion; Supplementary Fig. 1). The beginning of the isotope excursion occurs within an approximately 10-m-thick sequence of crevasse splay deposits; the lowest values occur with the first laterally continuous sand-body, demarcating the onset of the deposition of the Molina member. The excursion persists for an additional 30 m or so of the Molina member (Fig. 2; Supplementary Fig. 1; Supplementary Data). The Molina member can be continuously traced in outcrop for about 40 km both east–west and north–south in the study area around DeBeque, Colorado[18]. Equivalent sand-rich PETM intervals crop out about 90 km north and about 50 km east of DeBeque, and are recognizable in well data in intervening areas[16,18].

The Atwell Gulch member is a mud-dominated succession of purple, orange and red palaeosols (Supplementary Figs 2 and 3). Fluvial sand-bodies are thin and laterally restricted (Fig. 3a, b). Estimates of bank-full flow depths and widths indicate that the rivers that deposited these sand-bodies were relatively shallow and narrow, and that the bedforms within channel-fills were dominated by trough cross-bedding (Fig. 3c–e). Levee complexes are commonly associated with these sand-bodies, but crevasse splay deposits are rare (Supplementary Figs 2 and 3). In contrast, fluvial sand-bodies are thick, laterally continuous, and sheet-like within the Molina member (Fig. 3a, b). Bank-full flow depths are deeper, channels wider, and both display a greater range than in the other members (Fig. 3c, d). Upper-plane bed laminations are the dominant bedform observed, with palaeosols typically purple in colour, levee complexes absent, and crevasse splay deposits ubiquitous (Fig. 3e; Supplementary Figs 2–4).

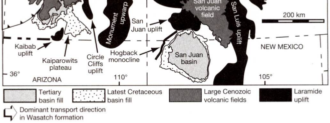


Figure 1 | Generalized geologic map showing major Laramide structures and associated basins. The Uinta and Piceance Creek basins were separate during the Palaeocene and the earliest Eocene epochs, and Cenozoic volcanic fields substantially post-date the deposition of the Wasatch formation.

---

[1] University of Wyoming, Department of Geology and Geophysics, 1000 East University Avenue, Laramie, Wyoming 82071, USA.

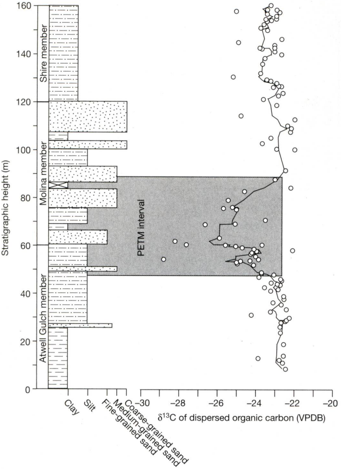


Figure 2 | Stratigraphic section through the middle portion of the Wasatch formation east of the town of DeBeque in Colorado, and the $\delta^{13}C$ record from dispersed organic carbon. The black line shows the five-point running average, and the width of individual data points represents analytical precision (about 0.1‰). Replicate analyses are ±0.3‰ of average values at a given stratigraphic height.

The strata of the overlying Shire member are similar to the Atwell Gulch member, except that the palaeosols are dominantly red and pink, and sand-bodies tend to be slightly thicker (Fig. 3; Supplementary Figs 2 and 3).

The sand-body and channel measurements are non-normally distributed based on Lilliefors statistical tests at the $\alpha = 0.05$ level. Non-parametric Kruskal–Wallis tests reject the null hypothesis that Molina member sand-body thickness (degrees of freedom, d.f. = 2; $\chi^2 = 46.1982$; $P \ll 0.001$), sand-body width (d.f. = 2; $\chi^2 = 35.1015$; $P \ll 0.001$), bank-full flow depth (d.f. = 2; $\chi^2 = 6.0541$; $P = 0.0485$), and flow width (d.f. = 2; $\chi^2 = 9.5330$; $P = 0.0085$) have the same median values as other members (Supplementary Data).

Previous authors ascribed the abrupt sedimentologic transition of the Molina member to either unroofing of Permian/Jurassic-aged aeolianites in the hinterland[18] or to a tectonically induced increase in sediment supply generated from surrounding Laramide uplifts[16] (Fig. 1). Sea level influences were probably unimportant given that western Colorado was separated from the palaeo-shoreline by around 1,000 km and several mountain ranges[14,19].

The unroofing hypothesis argues for a coarsening of the grain-size probability density function of flux from eroding mountain catchments[18]. We tested the hypothesis by assessing provenance changes within the Wasatch formation using U–Pb detrital zircon age spectra and their similarity to age spectra of the Glen Canyon group aeolianites[20], the proposed source for the Molina member[18]. All the major peaks and cumulative curves of the age spectra are nearly identical among the three members, indicating no major changes in sediment source during deposition (Fig. 4a; Supplementary Discussion). None show peak ages characteristic of older aeolianite sources from the Colorado plateau and the Uncompahgre uplift[20] (Fig. 4a). Furthermore, sandstone compositions are lithic arenites in all three members, unlike the quartz arenite composition of the aeolianites[21].

The tectonic hypothesis invokes simultaneous uplift events and corresponding increases in flux from the Uncompahgre and White River Laramide structures flanking the southwest and east of the study area, respectively[16] (Fig. 1). Constant subsidence rates (though limited in resolution) in the basin[22] and uniform palaeocurrent dispersal patterns (Fig. 4b) suggest that there were no periods of new or renewed rapid surface uplift and attendant flexural loading of the basin[23], nor deflection of rivers as a result of increased sediment flux from these margins. Moreover, the Molina member thins towards the proposed source areas[16], the opposite geometry of syntectonic deposits, which thicken towards the sediment source[23]. Alternatively, the Molina member could record a period of slowed subsidence with constant sediment flux, in which reduced accommodation causes progradation and denser amalgamation of coarse facies[23]. Existing subsidence histories[22] do not support this scenario either, but owing to age uncertainties neither tectonic scenario can be definitively disproved. However, we find them less parsimonious given the Molina member's correlation with the PETM and its sedimentologic coherency with proposed climatic changes.

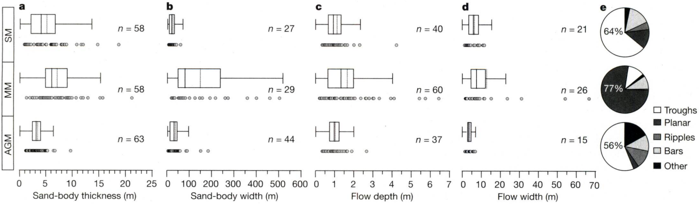


Figure 3 | Box and whisker plots of fluvial data from the Atwell Gulch, Molina and Shire members. AGM, Atwell Gulch member; MM, Molina member; SM, Shire member. Edges of boxes denote bounding quartiles, the black vertical lines represent median values, the grey vertical lines represent mean values, and whiskers denote the lower fence and upper fence (that is, 1.5 times the interquartile range). Grey circles denote individual data points. **a**, Sand-body thickness. **b**, Sand-body width perpendicular to local flow direction. **c**, Bank-full flow depth. **d**, Channel flow width perpendicular to mean local flow direction. **e**, Relative abundance of different bedform structures and bar clinoforms within sand-bodies.

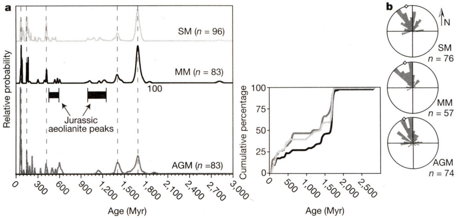


Figure 4 | Comparison of provenance and palaeodrainage patterns in the Atwell Gulch, Molina and Shire members. a, Normalized U–Pb age-distribution curves for detrital zircon populations (*n* is the number of grain ages determined). Key age peaks shared amongst the members are marked with dashed vertical lines. Age range of Jurassic aeolianite peaks are from ref. 20. Right, the cumulative distribution function incorporating errors. **b**, Rose diagrams showing uniform palaeocurrent direction throughout the Wasatch formation. Black diamonds represent vector means and *n* is the number of measurements.

Deeper and wider channels in the Molina member, assuming similar slopes, imply greater discharges and potentially higher mean annual precipitation. The greater range of values displayed in bankfull flow depths within the Molina member suggests larger variability in the channel-forming discharges, which may correlate with greater variability in the severity and intensity of rainfall events. Preservation of upper-flow-regime sedimentary structures, such as upper-plane bed laminations and climbing dunes (Supplementary Fig. 4), within channel deposits require either a peaked hydrograph or high in-channel sedimentation rates, or else they would have been reworked by waning flow stages[18,24]. Alternatively, such upper-flow-regime structures can occur in unusually shallow flows; however, this is difficult to reconcile with deeper flows indicated by bar clinoforms and would require sustained steepened river gradients to transport coarser sediment. The abundance of crevasse splay deposits and lack of well-developed levee complexes suggests that channel-breaching and flooding were common occurrences. Finally, a shift to purple palaeosols suggests less well drained conditions[25].

The high p_{CO_2} conditions of the PETM potentially instigated increased atmospheric humidity and intensified the hydrologic cycle[26]. Within the western interior of the USA, circulation models suggest the increased importance of convective atmospheric circulation off the palaeo-Gulf of Mexico, leading to enhanced monsoons[7,8]. Larger channels and preservation of upper-flow-regime structures are broadly consistent with the hypothesis, possibly with periods of higher runoff and greater channel-forming discharges associated with summer monsoonal rains. Yet increases in mean annual precipitation are more uncertain because channel-forming discharge will not necessarily reflect mean flow conditions. Overall, the monsoon hypothesis needs further testing at sites in other Laramide basins, especially since proxies in the Bighorn basin of Wyoming, around 500 km north, suggest drying trends during the PETM[6,25].

Numerical models predict extensive progradation of coarse-grained lithofacies due to increased discharge and sediment flux, reducing the rate of down-stream fining via selective deposition of coarser grains[1,3,4]. Greater rainfall leads to increased discharge (seasonal or otherwise), causing higher diffusivity and the capacity to transport coarse sediment in rivers[1,3]. Additional sediment flux to rivers will also cause progradation[1]. Continental-scale overturn of vegetation regimes during the PETM[5,6] undoubtedly remobilized sediment from basin hillslopes and floodplains. Similarly, greater hinterland catchment efflux is predicted during periods of vegetation overturn, precipitation increases and heightened storm intensity that together act to cleanse catchments of colluvium, enhance bedrock erosion by expanding drainage channel networks, increase bedrock channel incision rates, and accelerate sediment provision via landslides and other threshold-dependent hillslope processes[2,4].

Assuming these geomorphic processes led to increased sediment flux during the PETM, coupled with constant basin subsidence[22], the observed stratigraphic pattern implies the preferential bypass of fine-grained sediment through the basin towards the north (Fig. 1). The shift to sand-rich deposits is a consequence of selective deposition by alluvial rivers filling the basin[1,3]. Analogously, during the PETM in the Tremp-Graus basin of Spain, a vast conglomeratic braid-plain prograded owing to enhanced seasonal precipitation[27] with correlative shelf and bathyal marine sediments recording greater terrestrially derived clay accumulation[28]. Indeed, many marginal marine settings around the world record an increase in terrestrial clay deposition during the PETM[5,11,12]. The well-studied Bighorn basin may record the reverse scenario, in which basin-wide, enhanced palaeosol formation during the PETM[25] reflects reduced sediment supply due to lower diffusivity of basin rivers and catchment efflux brought on by decreases in precipitation by up to 40% (refs 6 and 25).

While increases in dissolved loads[10] and clay export[5,11,12] to oceans are restricted to the PETM interval, which implies fast response times in step with the climate change, the fluvial response in western Colorado persists 30 m beyond the isotope excursion (Fig. 2). Hysteresis effects such as this result from the dynamic coupling of hinterland erosional and basin depositional regimes and relate to how the perturbation is propagated (for example, a kinematic wave versus diffusion, respectively), the length scales they are transmitted over, and relaxation times for the reattainment of 'equilibrium' slope conditions[1–4]. Even in simplified two-dimensional models these effects may combine to maintain perturbed states for around 500 kyr (ref. 3). The likelihood of nonlinear sedimentation rates and imprecise age control precludes an accurate estimate for the relaxation time in western Colorado, though conservatively we suggest it was on the timescale of 10^5 years. More importantly, we emphasize the overall coherency in western Colorado with simplified and scaled-down model and experimental predictions, and that we expect future studies to find similar responses in other terrestrial sequences. If the pattern bears out, the high p_{CO_2} concentrations of the PETM did not only have far-reaching consequences for the evolution and ecology of biotic systems[5], but may also represent a global-scale 'clearing event' for geomorphic systems[13].

METHODS SUMMARY

U–Pb detrital zircon ages. Medium-grained sandstone samples were obtained near the base of fluvial sand-bodies (see Supplementary Discussion for stratigraphic positions). Samples were disaggregated, and zircons removed using standard water table, magnetic and heavy-liquid separation techniques. U–Pb determinations (about 100 unknown and about 35 standards per sample) were performed at the University of Arizona's LaserChron facility using a laser-ablation multicollector inductively coupled plasma mass spectrometer, measurement error is around 1–2% (2-sigma level). A 10% discordance filter was applied to the generated ages. See ref. 29 for detailed description of analytical methods.

Carbon isotope analyses. Samples for isotopic analysis were obtained by trenching until fresh rock was exposed. Approximately 40 mg of powdered sample were loaded into glass vials and loaded into a dry bath held at 50 °C. Then, 100 μl of 6 N HCl was added incrementally each day to the samples over the course of three days. Dried samples were weighed into tin capsules before introduction to the Thermal Finnigan Delta Plus XP elemental-analyser isotopic-ratio mass spectrometer housed at the University of Wyoming Stable Isotope Facility. Results are reported in δ notation with reference standard VPDB. Analytical error is around 0.1‰, and replicate analyses are ±0.3‰ of average values at a given stratigraphic height.

Stratigraphic data. Over 175 sand-bodies were examined during the course of this study across about 1,200 km^2. Fluvial data were obtained using a Jacob's staff and laser range finder. The accuracy of the laser range finder is ±0.1 m, and its precision is better than ±0.2 m. Flow depths were determined from relief on bar clinoforms and mud plugs. Flow widths were estimated from 1.5 times the toe-to-crest horizontal distance of bar clinoforms and mud-plug widths[30], corrected for local palaeocurrent direction relative to outcrop orientation. These estimates should be viewed as minimums because bar clinoform deposition may be oblique to the flow direction.

Received 7 February; accepted 14 August 2012.

Published online 24 October 2012.

1. Paola, C., Heller, P. L. & Angevine, C. L. The large-scale dynamics of grain-size variation in alluvial basins, 1: Theory. *Basin Res.* **4,** 73–90 (1992).
2. Tucker, G. E. & Slingerland, R. Drainage basin responses to climate change. *Wat. Resour. Res.* **33,** 2031–2047 (1997).
3. Armitage, J. J., Duller, R. A., Whittaker, A. C. & Allen, P. A. Transformation of tectonic and climatic signals from source to sedimentary archive. *Nature Geosci.* **4,** 231–235 (2011).
4. Allen, P. A. in *Landscape Evolution: Denudation, Climate and Tectonics Over Different Time and Space Scales* Vol. 269 (eds Gallacher, K., Jones, K. S. J. & Wainwright, J.) 7–28 (Geological Society Special Publications, 2008).
5. McInerney, F. A. & Wing, S. L. The Paleocene-Eocene Thermal Maximum: a perturbation of carbon cycle, climate, and biosphere with implications for the future. *Annu. Rev. Earth Planet. Sci.* **39,** 489–516 (2011).
6. Wing, S. L. *et al.* Transient floral change and rapid global warming at the Paleocene-Eocene boundary. *Science* **310,** 993–996 (2005).
7. Shellito, C. J., Sloan, L. C. & Huber, M. Climate model sensitivity to atmospheric CO$_2$ levels in the early-middle Paleogene. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **193,** 113–123 (2003).
8. Winguth, A., Shellito, C., Shields, C. & Winguth, C. Climate response at the Paleocene Eocene Thermal Maximum to greenhouse gas forcing—a model study with CCSM3. *J. Clim.* **23,** 2562–2584 (2010).
9. Murphy, B. H., Farley, K. A. & Zachos, J. C. An extraterrestrial ^{3}He-based timescale for the Paleocene-Eocene Thermal Maximum (PETM) from Walvis Ridge, IODP Site 1266. *Geochim. Cosmochim. Acta* **74,** 5098–5108 (2010).
10. Ravizza, G., Norris, R. N., Blusztajn, J. & Aubry, M.-P. An osmium isotope excursion associated with the late Paleocene thermal maximum: evidence of intensified chemical weathering. *Paleoceanography* **16,** 155–163 (2001).
11. Robert, C. & Kennett, J. P. Antarctic subtropical humid episode at the Paleocene–Eocene boundary—clay mineral evidence. *Geology* **22,** 211–214 (1994).
12. Gibson, T. G., Bybell, L. M. & Mason, D. B. Stratigraphic and climatic implications of clay mineral changes around the Paleocene/Eocene boundary of the northeastern US margin. *Sedim. Geol.* **134,** 65–92 (2000).
13. Jerolmack, D. J. & Paola, C. Shredding of environmental signals by sediment transport. *Geophys. Res. Lett.* **37,** L19401 (2010).
14. Dickinson, W. R. *et al.* Paleogeographic and paleotectonic setting of Laramide sedimentary basins in the central Rocky Mountain region. *Geol. Soc. Am. Bull.* **100,** 1023–1039 (1988).
15. Johnson, R. C. & May, F. Preliminary stratigraphic studies of the upper part of the Mesaverde Group, the Wasatch Formation, and the lower part of the Green River Formation, DeBeque area, Colorado, including environments of deposition and investigation of palynomorph assemblages. *USGS Misc. Field Investig.* Map MF-1050 (1978).
16. Johnson, R. C. & Flores, R. M. in *Piceance Basin 2003 Guidebook* (eds Peterson, K. M., Olson, T. M. & Anderson, D. A.) 21–61 (RMAG, 2003).
17. Burger, B. J. & Honey, J. G. Plesiadapidae (Mammalia, Primates) from the late Paleocene Fort Union Formation of the Piceance Creek Basin, Colorado. *J. Vert. Paleo.* **28,** 816–825 (2008).
18. Lorenz, J. C. & Nadon, G. C. Braided-river deposits in a muddy depositional setting: the Molina Member of the Wasatch Formation (Paleogene) west-central Colorado, U.S.A. *J. Sedim. Res.* **72,** 376–385 (2002).
19. Lillegraven, J. A. & Ostresh, L. M., Jr. in *Dawn of the Age of Mammals in the northern part of the Rocky Mountain Interior, North America* Vol. 243 (eds Bown, T. M. & Rose, K. D.) 1–30 (Geological Society of America, GSA Special Paper, 1990).
20. Dickinson, W. R. & Gehrels, G. E. U-Pb ages of detrital zircons from Permian and Jurassic eolian sandstones of the Colorado Plateau, USA: paleogeographic implications. *Sedim. Geol.* **163,** 29–66 (2003).
21. Blecha, A. M. & Gardner, M. H. Quantifying tectonic and climatic controls on alluvial architecture: Wasatch Formation of western Colorado. *GSA Abstr. Progr.* **36,** 307 (2004).
22. Johnson, S. Y. Phanerozoic evolution of sedimentary basins in the Uinta-Piceance Basin region, northwestern Colorado and northeastern Utah. *USGS Bull.* **1787-FF,** FF1–FF38 (1992).
23. Heller, P. L., Angevine, C. L., Winslow, N. S. & Paola, C. Two-phase stratigraphic model of foreland-basin sequences. *Geology* **16,** 501–504 (1988).
24. Fielding, C. R. Upper flow regimes sheets, lenses and scour fills: extending the range of architectural elements for fluvial sediment bodies. *Sedim. Geol.* **190,** 227–240 (2006).
25. Kraus, M. J. & Riggins, S. Transient drying during the Paleocene-Eocene Thermal Maximum (PETM): analysis of paleosols in the Bighorn Basin, Wyoming. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **245,** 444–461 (2007).
26. Bowen, G. J., Beerling, D. J., Koch, P. L., Zachos, J. C. & Quattlebaum, T. A humid climate state during the Palaeocene/Eocene thermal maximum. *Nature* **432,** 495–499 (2004).
27. Schmitz, B. & Pujalte, V. Abrupt increase in seasonal extreme precipitation at the Paleocene-Eocene boundary. *Geology* **35,** 215–218 (2007).
28. Schmitz, B. & Pujalte, V. Sea-level, humidity, and land-erosion records across the initial Eocene thermal maximum from a continental-marine transect in northern Spain. *Geology* **31,** 689–692 (2003).
29. Gehrels, G. E., Valencia, V. A. & Ruiz, J. Enhanced precision, accuracy, efficiency and spatial resolution of U-Th-Pb ages by LA-MC-ICPMS. *Geochem. Geophys. Geosyst.* **9** (2008).
30. Ethridge, F. G. & Schumm, S. A. in *Fluvial Sedimentology* (ed. Miall, A. D.) 703–721 (Canadian Society of Petroleum Geologists, 1978).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank personnel at the University of Arizona LaserChron Facility and University of Wyoming Stable Isotope Facility for assistance in analysing detrital zircon and carbon isotope samples. The International Association of Sedimentologists, the Tobacco Root Geological Society, the Colorado Scientific Society, and the Chevron Energy Technology Company provided funding for analyses and fieldwork. The research was completed while on a NSF Graduate Research Fellowship and Wyoming NASA Space Grant Consortium Fellowship to B.Z.F.

Author Contributions B.Z.F. and P.L.H. designed the study. B.Z.F. collected field data sets. B.Z.F. and M.T.C. carried out isotopic analyses. B.Z.F. wrote the manuscript. All authors contributed to the interpretations and conclusions presented.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.Z.F. (bforema1@uwyo.edu).

doi:10.1038/nature11565

The elusive Hadean enriched reservoir revealed by ^{142}Nd deficits in Isua Archaean rocks

Hanika Rizo[1], Maud Boyet[1], Janne Blichert-Toft[2], Jonathan O'Neil[1], Minik T. Rosing[3] & Jean-Louis Paquette[1]

The first indisputable evidence for very early differentiation of the silicate Earth came from the extinct ^{146}Sm–^{142}Nd chronometer. ^{142}Nd excesses measured in 3.7-billion-year (Gyr)-old rocks from Isua[1,2] (southwest Greenland) relative to modern terrestrial samples imply their derivation from a depleted mantle formed in the Hadean eon (about 4,570–4,000 Gyr ago). As dictated by mass balance, the differentiation event responsible for the formation of the Isua early-depleted reservoir must also have formed a complementary enriched component. However, considerable efforts to find early-enriched mantle components in Isua have so far been unsuccessful[3–7]. Here we show that the signature of the Hadean enriched reservoir, complementary to the depleted reservoir in Isua, is recorded in 3.4-Gyr-old mafic dykes intruding into the Early Archaean rocks. Five out of seven dykes carry ^{142}Nd deficits compared to the terrestrial Nd standard, with three samples yielding resolvable deficits down to −10.6 parts per million. The enriched component that we report here could have been a mantle reservoir that differentiated owing to the crystallization of a magma ocean, or could represent a mafic proto-crust that separated from the mantle more than 4.47 Gyr ago. Our results testify to the existence of an enriched component in the Hadean, and may suggest that the southwest Greenland mantle preserved early-formed heterogeneities until at least 3.4 Gyr ago.

^{142}Nd is produced by decay of short-lived ^{146}Sm (half-life 68 million years, Myr; ref. 8). Because the ^{146}Sm–^{142}Nd chronometer was present only during the first ~500 Myr of Solar System history, it is a powerful tool for understanding the evolution of the silicate Earth during the Hadean. Archaean rocks from the Amitsoq Complex (southwest Greenland) show small positive deviations in ^{142}Nd/^{144}Nd ratio relative to other terrestrial samples[1,2,9–13]. This isotopic signature implies their derivation from a depleted mantle formed in the Hadean, and therefore documents that the silicate Earth experienced a differentiation event during the first hundreds of Myr of Solar System history. The early differentiation event resulting in the formation of the depleted reservoir must also have formed a complementary enriched reservoir. Indirect evidence for the existence of this component comes from the independent observations that both the magnitude of ^{142}Nd excesses in Archaean rocks (Fig. 1) and the Lu/Hf and Sm/Nd ratios of komatiite sources decrease over time[14]. This may reflect the partial remixing of Hadean enriched and depleted reservoirs. Moreover, the presence of an enriched mafic crust is required to explain the Hf isotope compositions of the detrital Hadean Jack Hills zircons[15,16]. Additionally, an enriched basaltic crust may have been identified in the Nuvvuagittuq greenstone belt (Québec, Canada). Mafic samples from this belt have ^{142}Nd/^{144}Nd values lower than the terrestrial standard[17], and the positive correlation between ^{142}Nd/^{144}Nd and ^{147}Sm/^{144}Nd

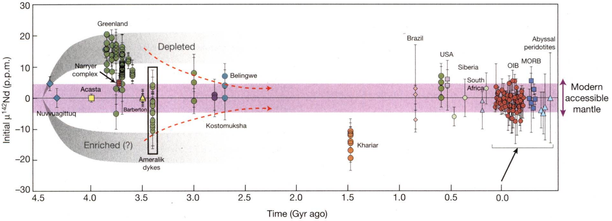


Figure 1 | Compilation of all published initial ^{142}Nd/^{144}Nd ratios for terrestrial samples. Data are taken from refs 2–7, 9, 12, 13 and 17, except for the data for the Ameralik dykes, which are from the present work (green symbols in black rectangle). Data are plotted as μ^{142}Nd, that is, deviations in p.p.m. relative to the terrestrial Nd standard JNdi-1. The purple shaded area represents the external analytical error of ±5 p.p.m. (2σ) as a function of the time of emplacement of each sample. The 3.7–3.8-Gyr-old Greenland samples have well-resolved positive ^{142}Nd anomalies, which suggests that their source was depleted in incompatible elements and formed in the Hadean (~4.47 Gyr ago[9]). The fractionation event responsible for the formation of the Greenland early-depleted mantle (grey shaded area labelled 'Depleted') must have created a complementary enriched reservoir (grey shaded area labelled 'Enriched(?)'). The only samples that seem to derive from a low-^{142}Nd/^{144}Nd source are the Khariar samples[4], but this data set still needs to be duplicated. Note that ^{142}Nd anomalies seem to decrease with time, which may reflect the remixing of both early-depleted and enriched reservoirs (red dashed curves). OIB (ocean island basalt), MORB (mid-ocean-ridge basalt) and abyssal peridotites are all modern samples (that is, 0 Gyr ago, as indicated by the bracket). Samples derived from the modern accessible mantle are indicated by the purple double-headed arrow.

---

[1]Clermont Université, Université Blaise Pascal, Laboratoire Magmas et Volcans, CNRS UMR 6524, IRD R 163, BP 10448, 63000 Clermont-Ferrand, France. [2]Laboratoire de Géologie de Lyon, Ecole Normale Supérieure de Lyon and Université Claude Bernard Lyon 1, CNRS UMR 5276, 69007 Lyon, France. [3]Natural History Museum of Denmark and Nordic Center for Earth Evolution, University of Copenhagen, DK-1350 København, Denmark.

ratios suggests that these rocks crystallized ~4.4 Gyr ago (calculated with the new ^{146}Sm decay constant[8]). All other efforts to identify early-enriched mantle components using the ^{146}Sm–^{142}Nd system have so far been either unsuccessful or yielded controversial results[3–7].

In a new quest for the elusive enriched Hadean reservoir, we analysed the mafic Ameralik dykes from the Amitsoq Complex for major- and trace-element abundances and ^{146,147}Sm–^{142,143}Nd, ^{176}Lu–^{176}Hf and U–Pb systematics. The Amitsoq Complex is located in the Nuuk region and includes the Isua Supracrustal Belt (ISB). Mafic dykes known as the 'Ameralik dykes' intrude into both the ISB and its enveloping gneisses. We analysed 12 metadoleritic dykes intruding into the ISB, as well as five samples that included margins and central cumulative parts of noritic dykes intruding into the Amitsoq gneisses (Supplementary Information).

The trace-element and isotopic compositions of the centres of the cumulate noritic dykes differ significantly from those of their margins, which have compositions identical to the metadolerite dykes. The cumulate noritic dykes have light rare-earth element (LREE)-enriched chondrite-normalized (CN) profiles ($La/Sm_{CN} = 3.3–3.6$) and low ^{147}Sm/^{144}Nd (0.1241–0.1247) and ^{176}Lu/^{176}Hf (0.0156–0.0175) ratios (Supplementary Information). These dykes therefore seem to have been contaminated during of after emplacement into their host gneisses. The best explanation for the contamination of the cumulate norites is via LREE-enriched fluid phases expelled from the surrounding gneisses (see Supplementary Information for more details). Furthermore, these contaminated samples (the noritic dykes) have resolvable ^{142}Nd excesses (Fig. 2) and ^{142}Nd/^{144}Nd ratios similar to the country rock (the host gneisses[12]). This suggests that the assimilation of fluids from the country rock has also modified the ^{142}Nd/^{144}Nd ratio of the noritic samples. These results demonstrate that great care must be exercised in the use of the ^{146}Sm–^{142}Nd system, as ^{142}Nd signatures can be inherited by contamination.

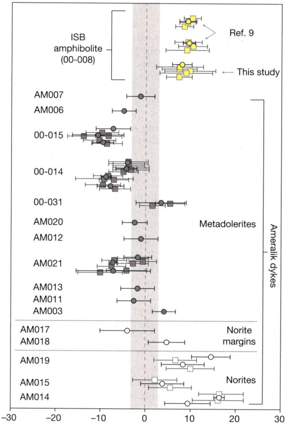


Figure 2 | μ^{142}Nd values measured for the Ameralik dykes. The ^{142}Nd/^{144}Nd data are expressed in p.p.m. deviations (μ^{142}Nd) relative to the JNdi-1 terrestrial standard. The shaded area represents the external analytical error of ±3 p.p.m. (2σ, $n = 50$). The same sample dissolution was analysed one to five times and the different runs are shown by squares. Circles represent separate sample dissolutions and the averages of the different runs of the same dissolution. Errors are 2σ. Sample 00-008 (3.7 Gyr-old ISB amphibolite) was analysed during the same analytical session and yielded a well-resolved ^{142}Nd excess of $+8.0 \pm 2.2$ p.p.m. (2σ), identical within error to the data obtained previously on the same sample[9]. Samples AM021, 00-014 and 00-015 all have well resolved negative anomalies down to −10.6 p.p.m., suggesting they have preserved the signature from an enriched Hadean source. All the ^{142}Nd results are reported in Supplementary Table 5.

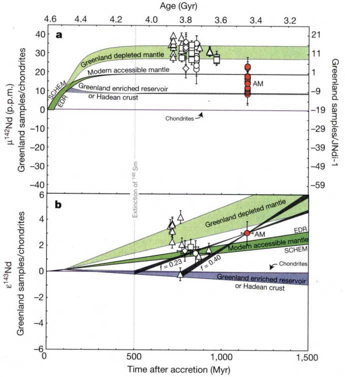


Figure 3 | Evolution model of the Ameralik dyke reservoir. a, b, The μ^{142}Nd (**a**) and ε^{143}Nd (**b**) values of southwest Greenland samples versus time after accretion (Myr, bottom scale) and versus age (Gyr, top scale). μ^{142}Nd and ε^{143}Nd errors are 2σ. **a**, The ^{142}Nd/^{144}Nd data for the Ameralik dykes (this study, filled symbols labelled 'AM') and for all other southwest Greenland samples (open symbols: circles[9], squares[11], triangles[12] and diamonds[13]), expressed in p.p.m. deviations relative to the chondritic value (Supplementary Information) and to the JNdi-1 terrestrial standard on the left and right scales, respectively. **b**, Initial ^{143}Nd/^{144}Nd calculated from isochrons for data from refs 9, 11 and 13, and for each sample for data from ref. 12, using $\lambda^{147}\mathrm{Sm} = 0.654 \times 10^{-11}$ and CHUR values from ref. 30. Initial values of ^{143}Nd/^{144}Nd are given in the ε^{143}Nd notation relative to chondrites[30]. The chemical fractionation between Sm and Nd is expressed as the fractionation factor $f_{Sm/Nd}$ (with $f_{Sm/Nd} = Sm/Nd_{reservoir}/Sm/Nd_{source}) - 1$; ref. 23). The Ameralik dykes show ^{142}Nd deficits compared to the terrestrial Nd standard and modern terrestrial rocks. These samples, therefore, have preserved the signature of an enriched Hadean reservoir, which differentiated at ~4.47 Gyr ago from the modern accessible mantle (both models SCHEM[21] or EDR[22] are represented, see Supplementary Information). The minimum enrichment value for the Hadean reservoir is μ^{142}Nd = −10 p.p.m. compared to JNdi-1 in order to explain the most enriched Ameralik samples. This early-enriched component satisfies the characteristics for the complementary early-depleted Isua reservoir. To explain simultaneously the positive ε^{143}Nd initial values of the Ameralik dykes, this component re-differentiated at ~770 Myr ago at the latest, for a maximum $f_{Sm/Nd}$ of +0.4. Shaded light green area represents the Greenland depleted mantle evolution. The dark green area represents the modern accessible mantle. The purple line represents the chondrites and the blue area represents the Greenland enriched reservoir or Hadean crust evolution.

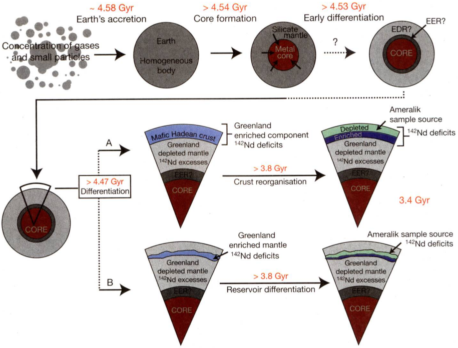

~ 4.58 Gyr
Earth's accretion
> 4.54 Gyr
Core formation
> 4.53 Gyr
Early differentiation
Concentration of gases and small particles
Earth
Homogeneous body
Silicate mantle
Metal core
?
EDR?
EER?
CORE
A
> 4.47 Gyr
Differentiation
B
CORE
Mafic Hadean crust
Greenland depleted mantle
^{142}Nd excesses
EER?
CORE
Greenland enriched component
^{142}Nd deficits
> 3.8 Gyr
Crust reorganisation
Ameralik sample source
Depleted
Enriched
^{142}Nd deficits
Greenland depleted mantle
^{142}Nd excesses
EER?
CORE
3.4 Gyr
Greenland enriched mantle
^{142}Nd deficits
Greenland depleted mantle
^{142}Nd excesses
EER?
CORE
> 3.8 Gyr
Reservoir differentiation
Ameralik sample source
^{142}Nd deficits
Greenland depleted mantle
^{142}Nd excesses
EER?
CORE

Figure 4 | From the accretion of the Earth to the differentiation of the Ameralik dyke source. The Earth accreted from gases and particles ~4.58 Gyr ago. The metallic core formed rapidly (>4.54 Gyr ago) and separated from the silicates. Comparison with chondrites suggests that the Earth evolved since the beginning of the Solar System history (or a few tens of million years after) with a high Sm/Nd ratio. Therefore, Earth's primitive mantle could have undergone Sm/Nd differentiation >4.53 Gyr ago leading to the formation of an early depleted reservoir (EDR) and a complementary early enriched reservoir (EER). The primitive mantle then underwent differentiation >4.47 Gyr ago, producing the Greenland depleted mantle. The same differentiation event led to a complementary enriched reservoir. The Ameralik dykes (southwest Greenland) recorded this enriched component, which must have differentiated before the extinction of ^{146}Sm (which produced ^{142}Nd) and, hence, before ~500 Myr ago. The most likely scenario is that this enriched component differentiated at >4.47 Gyr ago and that it constitutes the complementary part to Greenland's depleted reservoir. The enriched component could have been either: a Hadean mafic protocrust (path A on the figure) or a mantle reservoir (path B). This component re-differentiated at the latest 3.8 Gyr ago.

The ^{147}Sm–^{143}Nd whole-rock isochron age obtained for the metadolerites and the margins of the noritic dykes (3,403 ± 250 Myr ago) is indistinguishable from the mean ^{207}Pb/^{206}Pb age recorded by zircons from the norites (3,421 ± 34 Myr ago; Supplementary Information). This age is also consistent with other noritic dyke age determinations[18–20] and confirms that the Ameralik dyke swarm was emplaced at ~3.4 Gyr ago. The initial ε^{143}Nd of $+3.0 \pm 0.9$ (2σ, $n = 14$) from the ^{147}Sm–^{143}Nd regression (mean squared weighted deviation MSWD = 2.4) indicates that the source of these dykes was suprachondritic. (Here $\varepsilon^{143}\mathrm{Nd}(t) = \{[(^{143}\mathrm{Nd}/^{144}\mathrm{Nd})_{\mathrm{sample}}(t)/(^{143}\mathrm{Nd}/^{144}\mathrm{Nd})_{\mathrm{CHUR}}(t)] - 1\} \times 10^4$, where t is the time of dyke emplacement.) The dyke swarm thus derived from a source depleted in incompatible elements and characterized by a high Sm/Nd ratio. Similar constraints could not be acquired with the Lu–Hf isotope system because the samples scatter significantly in ^{176}Lu/^{177}Hf–^{176}Hf/^{177}Hf space (MSWD = 21; Supplementary Information). This scatter, observed in Lu–Hf but not in Sm–Nd isotope space, could be due to Lu^{3+} being a rare earth element (REE), whereas Hf^{4+} is a high-field strength element, making it possible to fractionate Lu from Hf. In contrast, Sm^{3+} and Nd^{3+} are neighbouring REEs, making it particularly difficult to disturb these two elements independently of each other. Importantly, except for two samples, all the metadolerite dykes have ^{142}Nd deficits compared to the terrestrial JNdi-1 Nd standard, with three samples showing resolvable ^{142}Nd deficits down to −10.6 p.p.m. (Fig. 2). These results were duplicated up to four times using different sample dissolution and digestion techniques (see Methods for more details).

The negative μ^{142}Nd values ($\mu^{142}\mathrm{Nd} = \{[(^{142}\mathrm{Nd}/^{144}\mathrm{Nd})_{\mathrm{sample}}/(^{142}\mathrm{Nd}/^{144}\mathrm{Nd})_{\mathrm{chondrites\ or\ JNdi\text{-}1}}] - 1\} \times 10^6$) of the Ameralik dykes are consistent with their derivation from an early-enriched component; however, their positive $\varepsilon^{143}\mathrm{Nd}_{3.4\mathrm{Gyr}}$ value (+3.0) indicates that the source of these samples has experienced a multi-stage history. Any Sm/Nd fractionation event occurring after the complete decay of ^{146}Sm will affect the ^{147}Sm–^{143}Nd system exclusively, hence resulting in the decoupling of the long-lived and short-lived Sm–Nd chronometers. Several models were tested, and the most likely scenario is presented in Fig. 3. This model is based on two assumptions. First, the enriched component identified in the source of the Ameralik dykes was formed contemporaneously with the depleted Isua reservoir (>4.47 Gyr ago[9,11]) and thus most probably is related to the crystallization of an early terrestrial magma ocean. This assumption is further consistent with the close geographic association of the Ameralik dykes and the ISB rocks. Second, since the beginning of Solar System history (or a few tens of million years after) the terrestrial mantle evolved with a Sm/Nd ratio higher than that of chondrites. Such an evolution has been proposed to explain the difference in ^{142}Nd composition between chondrites and terrestrial samples[21,22].

The lowest ^{142}Nd values in the Ameralik dykes are consistent with the formation of an enriched component with an ^{147}Sm/^{144}Nd ratio of 0.19. The formation of a reservoir with a higher ^{147}Sm/^{144}Nd ratio would produce ^{142}Nd deficits too small for what is observed in the Ameralik dykes. The positive initial $\varepsilon^{143}\mathrm{Nd}_{3.4\mathrm{Gyr}}$ of the dykes indicates that their source underwent a second Sm/Nd differentiation event to produce a depleted reservoir. This later event occurred after the extinction of ^{146}Sm, thus preserving the ^{142}Nd deficits while the ^{143}Nd isotopic compositions evolved with time to a positive ε^{143}Nd value. This differentiation event therefore could not have taken place earlier than ~500 Myr after Earth's accretion. However, the later this differentiation event occurred, the more depleted the newly formed reservoir must have been to evolve to ε^{143}Nd of +3.0 by 3.4 Gyr ago. The latest plausible time of re-differentiation is thus controlled by the maximum $f_{\mathrm{Sm/Nd}}$ possible, <0.40 (ref. 25) ($f_{\mathrm{Sm/Nd}}$ is defined[23] as $(\mathrm{Sm/Nd}_{\mathrm{reservoir}}/\mathrm{Sm/Nd}_{\mathrm{source}}) - 1$); Fig. 3). Using this maximum $f_{\mathrm{Sm/Nd}}$ value, the latest age of re-differentiation is constrained to ~670 Myr after the formation of the early-enriched component (Fig. 3). Depletion of an earlier enriched component by secondary processes has been previously proposed[24] in a scenario where a thick basaltic Hadean proto-crust (the enriched reservoir) would be thermally unstable, leading to partial melting at its base and leaving a depleted residue.

Partial melting after crystallization of a magma ocean could have formed the first mafic Hadean crust. However, our data cannot be accounted for if the crust formed after 4.3 Gyr ago. In such a scenario, the $f_{\mathrm{Sm/Nd}}$ needed to explain the lowest ^{142}Nd/^{144}Nd ratios measured in the Ameralik dykes is -0.46, which is even more enriched than typical granitoids ($f_{\mathrm{Sm/Nd}} = -0.40$; ref. 23) (Supplementary Information). Alternatively, the Earth's primitive mantle could have undergone Sm/Nd differentiation >4.53 Gyr ago leading to the formation of an early depleted reservoir (EDR) and a complementary early enriched reservoir (EER)[21] (Fig. 4). If the Earth started out with a chondritic ^{142}Nd composition, mass balance constraints would require the EER to have μ^{142}Nd in the range of -38 to -54 p.p.m. (ref. 26). The lack of samples with such characteristics indicates that the EER has never participated in crustal magmatism. Could the Ameralik sample source have differentiated from the EER? A depleted reservoir separating from an EER-like source before the extinction of ^{146}Sm would evolve towards less negative μ^{142}Nd values than the EER and positive initial ε^{143}Nd values. This could fit the source for the Ameralik dykes. In order to produce a $\sim$10 p.p.m. ^{142}Nd deficit at 3.4 Gyr ago, the latest possible time for the differentiation is $\sim$100 Myr after the first solids formed in the early Solar System (see Supplementary Information for more details). Although this scenario adequately explains the present data, it is highly unlikely that the Ameralik dykes sampled the EER. The ^{142}Nd excess measured in all modern terrestrial samples relative to chondrites requires that the EER has not been remixed in the mantle. Moreover, the EER is probably located in the deep mantle, because mass balance calculations predict a reservoir larger than the size of the continental crust[26].

The early-enriched component identified in this study satisfies the characteristics of the complementary early-depleted Isua reservoir. This early-enriched component has been missing since the first ^{142}Nd results[10]. Our findings attest to the existence of either an enriched mantle reservoir or a proto-crust in the Hadean (Fig. 4). It is difficult to argue which of these two protoliths is the best candidate for the source of the Ameralik dykes. However, if the Ameralik source was an enriched mantle reservoir, it demonstrates that the southwest Greenland mantle preserved early-formed heterogeneities until at least 3.4 Gyr ago. This would imply re-mixing of heterogeneities at rates of the order of 1 Gyr, which are 10 times longer than previously estimated[27,28], but in agreement with recent ^{182}W findings[29]. Conversely, if the Ameralik source represents proto-crust isolated from mantle convection, it no longer places constraints on the secular survival of mantle heterogeneities, but instead suggests that the Early Archaean crust could preserve relicts formed in Hadean times.

METHODS SUMMARY

The whole-rock samples were powdered and dissolved in high-pressure, steel-jacketed Parr Teflon bombs. REE and Hf fractions were separated from the bulk samples using ion-exchange columns. For the ^{142}Nd analyses, Ce and most of the Sm, which together produce isobaric interferences on ^{142}Nd, ^{144}Nd, ^{148}Nd and ^{150}Nd, were eliminated through two consecutive passes on cation-exchange columns using 2-methylactic acid (0.2 M and pH 4.7) as medium. The rest of the Sm was removed on an Ln-Spec column. Total Nd blanks were <60 pg. The five samples showing resolvable ^{142}Nd anomalies were dissolved up to four times using different digestion techniques (Savillex vials and high-pressure, steel-jacketed Parr Teflon bombs) and gave identical ^{142}Nd results within the quoted errors. ^{142}Nd isotope measurements were done as Nd^+ using zone-refined double Re filaments. Isotopic analyses were performed using a dynamic routine where Faraday cups are centred successively on masses 145 and 143. The terrestrial Nd standard JNdi-1 was measured often, and the average long-term value for five analytical sessions gave 1.141840 ± 0.000003 (2σ, $n = 50$) for ^{142}Nd/^{144}Nd. The same dissolution of each Ameralik sample was analysed one to five times depending on the Nd concentrations. The external reproducibility on the ^{142}Nd/^{144}Nd ratio was, on average, better than 5 p.p.m., which is similar to the reproducibility obtained on repeated measurements of the terrestrial standard. ^{142}Ce and ^{144}Sm were never high enough to produce corrections higher than the precision obtained on the ^{142}Nd/^{144}Nd ratio. Moreover, the measured ^{142}Nd/^{144}Nd ratios are not correlated with the amounts of Ce or Sm. The Ameralik samples with ^{142}Nd deficits do not exhibit negative ^{148}Nd and ^{150}Nd anomalies, which corroborate that their anomalous Nd isotopic compositions are not a consequence of mixing of variably depleted domains on the filament.

Full Methods and any associated references are available in the online version of the paper.

Received 27 May; accepted 30 August 2012.

1. Boyet, M. *et al.* ^{142}Nd evidence for early Earth differentiation. *Earth Planet. Sci. Lett.* **214,** 427–442 (2003).
2. Caro, G., Bourdon, B., Birck, J. L. & Moorbath, S. High precision ^{142}Nd/^{144}Nd measurements in terrestrial rocks: constraints on the early differentiation of the Earth's mantle. *Geochim. Cosmochim. Acta* **70,** 164–191 (2006).
3. Andreasen, R., Sharma, M., Subbarao, K. V. & Viladkar, S. G. Where on Earth is the enriched Hadean reservoir? *Earth Planet. Sci. Lett.* **266,** 14–28 (2008).
4. Upadhyay, D., Scherer, E. & Mezger, K. ^{142}Nd evidence for an enriched Hadean reservoir in cratonic roots. *Nature* **459,** 1118–1121 (2009).
5. Murphy, D. T. *et al.* In search of a hidden long-term isolated sub-chondritic ^{142}Nd/^{144}Nd reservoir in the deep mantle: implications for the Nd isotope systematics of the Earth. *Geochim. Cosmochim. Acta* **74,** 738–750 (2010).
6. Cipriani, A., Bonatti, E. & Carlson, R. W. Non-chondritic ^{142}Nd in sub-oceanic mantle peridotites. *Geochem. Geophys. Geosyst.* **12,** 1–8 (2011).
7. Jackson, M. G. & Carlson, R. W. Homogeneous superchondritic ^{142}Nd/^{144}Nd in the mid-ocean ridge basalt and ocean island basalt mantle. *Geochem. Geophys. Geosyst.* **13,** 1–10 (2012).
8. Kinoshita, N. *et al.* A shorter ^{146}Sm half-life measured and implications for ^{146}Sm-^{142}Nd chronology in the Solar System. *Science* **335,** 1614–1617 (2012).
9. Rizo, H., Boyet, M., Blichert-Toft, J. & Rosing, M. Combined Nd and Hf isotope evidence for deep-seated source of Isua lavas. *Earth Planet. Sci. Lett.* **312,** 267–279 (2011).
10. Harper, C. L. & Jacobsen, S. B. Evidence from coupled ^{147}Sm–^{143}Nd and ^{146}Sm–^{142}Nd systematics for very early (4.5-Gyr) differentiation of the Earth's mantle. *Nature* **360,** 728–732 (1992).
11. Caro, G., Bourdon, B., Birck, J.-L. & Moorbath, S. ^{146}Sm–^{142}Nd evidence from Isua metamorphosed sediments for early differentiation of the Earth's mantle. *Nature* **423,** 428–432 (2003).
12. Bennett, V. C., Brandon, A. D. & Nutman, A. P. Coupled ^{142}Nd-^{143}Nd isotopic evidence for Hadean mantle dynamics. *Science* **318,** 1907–1910 (2007).
13. Boyet, M. & Carlson, R. W. A new geochemical model for the Earth's mantle inferred from ^{146}Sm-^{142}Nd systematics. *Earth Planet. Sci. Lett.* **250,** 254–268 (2006).
14. Blichert-Toft, J. & Puchtel, I. S. Depleted mantle sources through time: evidence from Lu-Hf and Sm-Nd isotope systematics of Archean komatiites. *Earth Planet. Sci. Lett.* **297,** 598–606 (2010).
15. Blichert-Toft, J. & Albarède, F. Hafnium isotopes in Jack Hills zircons and the formation of the Hadean crust. *Earth Planet. Sci. Lett.* **265,** 686–702 (2008).
16. Kemp, A. I. S. *et al.* Hadean crustal evolution revisited: new constraints for Pb-Hf isotope systematics of the Jack Hills zircons. *Earth Planet. Sci. Lett.* **296,** 45–56 (2010).
17. O'Neil, J., Carlson, R. W., Francis, D. & Stevenson, R. K. Neodymium-142 evidence for Hadean mafic crust. *Science* **321,** 1828–1831 (2008).
18. Gill, C. O. & Bridgwater, D. The Ameralik dykes of west Greenland, the earliest known basaltic rocks intruding stable continental crust. *Earth Planet. Sci. Lett.* **29,** 276–282 (1976).
19. Nielsen, S. G., Baker, A. J. & Krogstad, E. J. Petrogenesis of an early Archean (3.4 Ga) norite dyke, Isua, West Greenland: evidence for early Archean crustal recycling? *Precambr. Res.* **118,** 133–148 (2002).
20. Nutman, A. P., Friend, C. R. L., Bennett, V. C. & McGregor, V. R. Dating of Ameralik dike swarms of the Nuuk district, southern West Greenland: mafic intrusion events starting from 3510 Ma. *J. Geol. Soc. Lond.* **161,** 421–430 (2004).
21. Boyet, M. & Carlson, R. W. ^{142}Nd evidence for early (>4.53 Ga) global differentiation of the silicate Earth. *Science* **309,** 576–581 (2005).
22. Caro, G. & Bourdon, B. Non-chondritic Sm/Nd ratio in the terrestrial planets: consequences for the geochemical evolution for the mantle-crust system. *Geochim. Cosmochim. Acta* **74,** 3333–3349 (2010).
23. DePaolo, D. J. in *Neodymium Isotope Geochemistry. An Introduction* 36 (Minerals and Rocks Series no. 20, Springer, 1988).
24. Kamber, B. S., Collerson, K. D., Moorbath, S. & Whitehouse, M. J. Inheritance of early Archean Pb-isotope variability from long-lived Hadean protocrust. *Contrib. Mineral. Petrol.* **145,** 25–46 (2003).
25. Bennett, V. C., Nutman, A. P. & McCulloch, M. T. Nd isotopic evidence for transient, highly depleted mantle reservoirs in the early history of the Earth. *Earth Planet. Sci. Lett.* **119,** 299–317 (1993).
26. Carlson, R. W. & Boyet, M. Composition of the Earth's interior: the importance of early events. *Phil. Trans. R. Soc. A* **366,** 4077–4103 (2008).
27. Blichert-Toft, J. & Albarède, F. Short-lived chemical heterogeneities in the Archean mantle with implications for mantle convection. *Science* **263,** 1593–1596 (1994).
28. Coltice, C. & Schmalzl, J. Mixing time in the mantle of the early Earth derived from 2-D and 3-D numerical simulations of convection. *Geophys. Res. Lett.* **33,** L23304 (2006).

29. Touboul, M., Putchtel, I. S. & Walker, R. J. ^{182}W Evidence for long-term preservation of Early mantle differentiation products. *Science* **335,** 1065–1069 (2012).
30. Bouvier, A., Vervoort, J. D. & Patchett, P. J. The Lu-Hf and Sm-Nd isotopic composition of CHUR: constraints from unequilibrated chondrites and implications for the bulk composition of terrestrial planets. *Earth Planet. Sci. Lett.* **273,** 48–57 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank C. Bosq for providing clean-laboratory facilities, D. Auclair for assistance with the TIMS, P. Télouk for assistance with the MC-ICP-MS, and J. L. Devidal for assistance with the Microprobe and the LA-ICP-MS. M. Benbakkar carried out the major element analyses. We thank A. Brandon and M. Jackson for comments that helped clarify the manuscript. We thank the Geological Survey of Japan for providing the isotopic standard JNdi-1. The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (to M.B.), the French Programme National de Planétologie of the Institut National des Sciences de l'Univers and Centre National d'Etudes Spatiales, and the French Agence Nationale de la Recherche (grants BEGDy and M&Ms) (to J.B.T.), and the French embassy in Denmark (to M.R.). This is Laboratory of Excellence ClerVolc contribution no. 30.

Author Contributions Samples from the Ameralik dykes were collected by H.R., M.B., J.O'N. and M.T.R. U–Pb analyses were carried out by J.-L.P. Preparation of samples, dissolution, spike calculations, chemical separation of Sm, Nd, Lu, and Hf and isotopic analyses and modelling of data were carried out by J.B.-T., H.R. and M.B. Manuscript preparation was carried out by H.R., M.B., J.O'N. and J.B.-T., and all the authors contributed to discussing the results and their implications.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.R. (h.rizo@opgc.univ-bpclermont.fr).

doi:10.1038/nature11512

Fault healing promotes high-frequency earthquakes in laboratory experiments and on natural faults

Gregory C. McLaskey[1]†, Amanda M. Thomas[2], Steven D. Glaser[1] & Robert M. Nadeau[2]

Faults strengthen or heal with time in stationary contact[1,2], and this healing may be an essential ingredient for the generation of earthquakes[1–3]. In the laboratory, healing is thought to be the result of thermally activated mechanisms that weld together micrometre-sized asperity contacts on the fault surface, but the relationship between laboratory measures of fault healing and the seismically observable properties of earthquakes is at present not well defined. Here we report on laboratory experiments and seismological observations that show how the spectral properties of earthquakes vary as a function of fault healing time. In the laboratory, we find that increased healing causes a disproportionately large amount of high-frequency seismic radiation to be produced during fault rupture. We observe a similar connection between earthquake spectra and recurrence time for repeating earthquake sequences on natural faults. Healing rates depend on pressure, temperature[4] and mineralogy[1], so the connection between seismicity and healing may help to explain recent observations of large megathrust earthquakes which indicate that energetic, high-frequency seismic radiation originates from locations that are distinct from the geodetically inferred locations of large-amplitude fault slip[5–7].

In laboratory measurements, static-fault frictional strength, μ_s, is generally observed to increase linearly with the logarithm of time in stationary contact, t_{hold}, according to

$$\mu_s(t_{\mathrm{hold}}) = \alpha_s + \beta_s \log_{10}(t_{\mathrm{hold}}) \qquad (1)$$

where β_s is the healing rate and α_s is the fault strength at time $t_{\mathrm{hold}} = 1$ s (refs 1–4, 8). These measurements are used to derive rate- and state-dependent friction laws[2,3] that have provided insight into fault behaviour ranging from slow slip to dynamic rupture[2,8–10]. Healing rates have also been inferred from repeating earthquake sequences[10–12] (RESs). These are sets of events with nearly identical waveforms, locations and magnitudes, and are thought to represent the repeated rupture of a patch of fault that is slowly loaded by aseismic slip of the surrounding material. Here we consider the stick-slip behaviour of a laboratory fault as a proxy for such a fault patch and compare our results with observations of RESs on the San Andreas fault. In addition to measuring static friction, slip and stress drop, we record the stress waves emitted during the rupture of the laboratory fault, which we call laboratory earthquakes (LabEQs). This facilitates a link between friction properties observed in the laboratory and earthquakes produced on natural faults.

Fault healing is typically attributed to an increase in either the area or the strength of asperity contacts due to 'creep'[2,3]. Mechanisms may include stress-induced diffusion, dislocation motion, chemically aided slow crack growth, dissolution–precipitation processes and other thermally activated processes[2,3,13–15]. Although specific mechanisms may differ, the overall effects of healing are remarkably similar. Equation (1) is applicable to rocks[1], metals[16], plastics[4] and paper[17], which suggests that the mechanics of healing are not greatly dependent on specific chemical or physical properties, but rely on universally observed surface properties such as multiscale roughness. A better understanding of the relationship between fault healing and earthquake generation may be the key to understanding the physics of earthquakes[18].

We use test blocks composed of the glassy polymer poly(methyl methacrylate) (PMMA). This and similar glassy polymers are commonly used as model materials for fault rupture and friction studies[2,4,19,20]. Friction on PMMA–PMMA interfaces obeys equation (1) and is well modelled by the rate- and state-dependent friction laws[2,4]. Because of PMMA's low hardness and melting temperature (~160 °C), the behaviour of PMMA–PMMA interfaces at room temperature and modest stress levels (100 kPa) may be somewhat representative of the behaviour of rocks at depth[20]. The similarities and differences between plastic and rock may serve as important points of comparison when studying the range of friction properties expected in the brittle–ductile spectrum of crustal deformation behaviour.

Stick-slip experiments are conducted at room temperature (20 °C) and humidity (30% relative humidity) on a direct-shear apparatus consisting of a PMMA slider block (181 mm long, 60 mm wide, 17 mm high) and a larger PMMA base plate (450 mm long, 300 mm wide, 36 mm high) (Fig. 1a, inset). With the normal force, F_N, held constant, the shear force, F_S, is increased until the sample undergoes a series of stick-slip instabilities, denoted events. The recurrence time, t_r, defined as the time since the previous event, is computed for each pair of consecutive events in the sequence. (Despite subtle differences between them[21], we assume that $t_r = t_{\mathrm{hold}}$.) Each event produces a LabEQ, which is recorded with piezoelectric sensors attached to the PMMA base plate. The slider block slips 50–200 μm during each event. Some slow premonitory slip (~2 μm) is often detected 1–2 ms before rapid slip commences. We detect no slip between events (to the ~1-μm noise level). The duration of slip for each event is approximately constant (8 ms) and is probably controlled by the combined stiffness of the apparatus and samples rather than fault rupture properties.

Load point displacement, x_{LP}, is controlled by turning a fine-threaded screw that presses against the trailing edge of the slider block. When the load point velocity, $v_{\mathrm{LP}} = \mathrm{d}x_{\mathrm{LP}}/\mathrm{d}t$, is systematically increased or decreased, large variations in t_r can be achieved in a single experimental run, while other experimental variables (F_N, surface conditions and so on) are kept constant. Typical results are shown in Fig. 1. To isolate cumulative wear and loading-rate effects, experimental runs were conducted in pairs: one with increasing v_{LP} (Fig. 1a) and one with decreasing v_{LP} (Fig. 1b). For every event in each stick-slip sequence, we measure $F_{\max}$ and $F_{\min}$ (Fig. 1) and calculate the stress drop $\Delta\tau = (F_{\max} - F_{\min})/A$, where A is the nominal fault area (0.0109 m^2). These parameters are plotted against $\log_{10}(t_r)$ (Fig. 1b, inset, and Supplementary Figs 2–4). Slopes, β, and intercepts, α, of the best-fit lines are reported in Supplementary Table 1. All tests show results consistent with equation (1) and previous work[1,3,4,8]. The increase in $\Delta\tau$ with increasing t_r is due to both an increase in $F_{\max}$ and a decrease in $F_{\min}$ with $\log_{10}(t_r)$ (refs 8, 21). In all cases, healing rates, β, are slightly larger for runs with decreasing v_{LP} than for runs with increasing v_{LP}, indicating a dependence on loading rate or stress time history.

---

[1]Department of Civil and Environmental Engineering, University of California, Berkeley, California 94720, USA. [2]Department of Earth and Planetary Science, Berkeley Seismological Laboratory, University of California, Berkeley, California 94720, USA. †Present address: United States Geological Survey, 345 Middlefield Road, MS 977, Menlo Park, California 94025, USA.

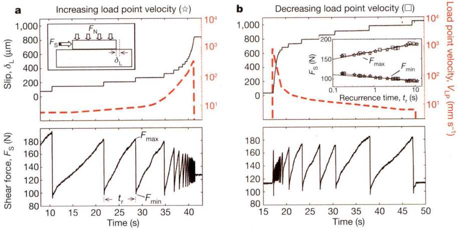


Figure 1 | Experimental data from a pair of healing tests. Shear force, F_S, load point velocity, v_{LP} (dashed line), and slip δ_L, measured from stick-slip experiments at $\sigma_n = 36$ kPa. All experiments were conducted in pairs, one with increasing v_{LP} (**a**) and one with decreasing v_{LP} (**b**). Insets: schematic of the apparatus (**a**) and the maximum (F_{max}) and minimum (F_{min}) shear forces measured for each event in the stick-slip sequence and plotted against the logarithm of t_r (**b**). Stars and squares are from runs with increasing and decreasing v_{LP}, respectively.

An example sequence of LabEQ seismograms is shown in Fig. 2. The interface properties, apparatus and specimen stiffness, sensor response and wave propagation characteristics do not change between successive events, so differences between LabEQs are attributed to variations in t_r. When each seismogram is scaled with respect to the total measured slip, δ_f, the low-frequency components (Fig. 2a) are nearly identical but the high-frequency components (Fig. 2b) depend strongly on t_r. Absolute source spectra were estimated for each LabEQ by removing the instrument and apparatus response functions from recorded signals by means of a ball-drop calibration source (Methods). Examples of absolute source displacement spectra are shown in Fig. 3a for three LabEQs from Fig. 2. Each source spectrum is roughly linear in log(ω), where ω is the angular frequency, so spectra were fitted with best-fit lines (not shown). Variations in the spectral slopes

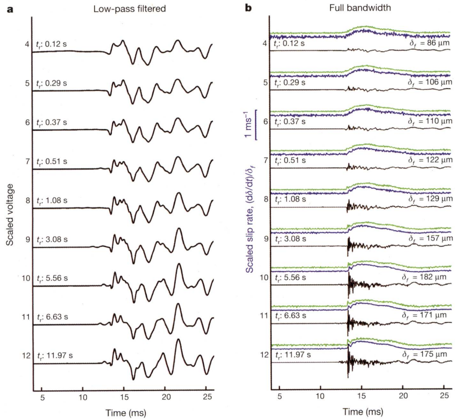


Figure 2 | Sequence of successive LabEQs. Events from an experimental run with decreasing v_{LP} (increasing time between successive events) at $\sigma_n = 130$ kPa and using a rough sample (run 45-R-Dec; see Supplementary Table 1). Each trace is scaled by the total measured slip, δ_f, indicated for each trace. **a**, Low-pass-filtered signals (1-kHz cut-off) illustrating the similarity of low-frequency waveforms. **b**, Full-bandwidth recorded LabEQ data (raw sensor output) plotted alongside scaled slip rates $(d\delta/dt)/\delta_f$, which are derived from slip measured at the leading (blue) and trailing (green) edges of the slider block and low-pass-filtered at 5 kHz to reduce high-frequency noise. The green curves have the same scale as the blue curves, and are offset for clarity.

of LabEQ source spectra are shown in Fig. 3b for all 46 events from four tests conducted at normal stress, $\sigma_n = 130\,\mathrm{kPa}$. These laboratory results show a disproportionate increase in high-frequency ground motions for greater t_r. Similar spectral changes were observed for all experiments, but are most pronounced for those conducted at higher σ_n. Peak high-frequency ground motions coincide with the initiation of slip, not maximum slip rate.

To complement the laboratory results, we analysed RESs on the San Andreas fault[22,23] that were perturbed by the 2004, moment-magnitude $M_w = 6$ Parkfield, California earthquake. As shown in Fig. 4, an increase in high frequencies with increasing t_r was observed for most RESs. Similar trends were found for the CA1 RES on the Calaveras fault[11]. If spectral changes were due to a propagation effect, such as damage from the Parkfield earthquake, we would expect the effects to be more pronounced in recordings from source–station ray paths that traverse long distances through zones of expected damage[24] (that is, near or within the fault zone and at shallower depths) (Supplementary Fig. 6 and Supplementary Table 2). Instead, many stations see similar spectral variations between the same events and spectral changes vary among RESs (Supplementary Fig. 1), so we suspect that spectral variations are dominantly controlled by changes in earthquake source characteristics and not path effects.

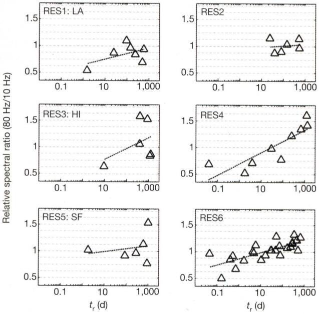


Figure 4 | Spectral changes of RESs near Parkfield. The SF, LA and HI RESs were targeted for penetration by the SAFOD deep-drilling experiment (sequences NW, SE and S1 in ref. 23). Relative spectral ratios are calculated from the ratio of relative spectral amplitudes at 75–85 Hz to those at 5–15 Hz (Supplementary Fig. 5 and Supplementary Table 3). Data points indicate the averages of the relative spectral ratios obtained from ground motions recorded at at least three stations for each event in each RES. Dotted lines show a linear best fit to the data, and a positive slope indicates increasing high-frequency ground motions (relative to low-frequency ground motion) with increasing $\log_{10}(t_r)$.

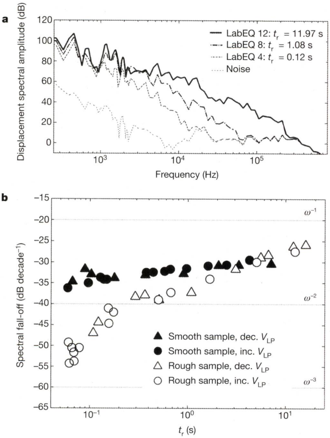


Figure 3 | LabEQ spectral changes with recurrence time. **a**, Source displacement spectra and the noise spectrum from three of the LabEQs shown in Fig. 2, which span two orders of magnitude in t_r. **b**, Slope of source spectrum as a function of recurrence time for all LabEQs from four experimental runs conducted at $\sigma_n = 130\,\mathrm{kPa}$. Only the frequency band with a signal-to-noise ratio greater than 6 dB was used for the calculation of these spectral slopes. For the rough sample, spectral slopes increase from $\omega^{-2.5}$ to $\omega^{-1.5}$ with increasing t_r. LabEQs generated from the smooth sample show subtle but systematic spectral changes.

Fault healing seems to cause spectral changes over a broad range of frequencies (Fig. 3a), so we propose that our observations are applicable not just to the small length scales and high frequencies of LabEQs, but to natural faults and great earthquakes as well. To discuss the underlying mechanisms of these spectral changes, we present a conceptual fault model in which both natural faults and those in the laboratory are composed of a large number of asperity contacts[2,25,26] with a distribution of strengths, which collectively sum to produce the static fault strength, μ_s. If the thermally activated healing mechanisms described above cause asperity contacts to strengthen at a rate proportional to the forces they support, then healing would promote a more heterogeneous spatial distribution of fault strength on the asperity scale. When this healed fault ruptures, its heterogeneous fault strength could cause perturbations in slip velocity that would generate high-frequency seismic waves[27]. Alternatively, if healing promotes a larger stress drop[8,21] or a more abrupt slip weakening behaviour, this would promote faster rupture propagation, which could also account for the enhanced high frequencies. This interpretation is consistent with a previous proposal that spectral changes observed for the CA1 RES signify shorter source duration, which could be explained by faster rupture propagation[11].

The spectral changes shown in Fig. 3b are somewhat analogous to those in Fig. 4, but when comparing the spectra of LabEQs with those from RESs, differences in rise time (the time during which a single point on the fault slips seismically) and rupture duration relative to the recorded frequency band should be taken into account. The LabEQ spectra shown in Fig. 3 are probably controlled by details of rupture propagation. Although the sample geometry and the resolution of the slip sensors do not permit a detailed analysis of dynamic rupture, Fig. 2b does show that slip accelerated more rapidly for events that healed longer. In the case of RESs, even the highest frequencies available for analysis (75–85 Hz) may be too low to contain much information about rupture propagation. Additionally, complicated interaction between rapid, unstable failure of the fault patch and stable slip imposed by slow slip of the surrounding region[9,10] may contribute to added differences between RESs at Parkfield and current laboratory analogues.

Dense seismic arrays have facilitated back-projection studies of recent megathrust earthquakes that highlight the temporal and spatial complexity of high-frequency seismic radiation and show that sources of high-frequency seismic waves are not spatially correlated with locations of maximum inferred fault slip[5–7]. A mechanism related to fault healing may be responsible for these puzzling observations, particularly for the March 2011 Tohoku earthquake, where high frequencies originated from deeper sections of the fault and contributed to strong ground accelerations felt in eastern Japan. Laboratory experiments on glassy polymers show that healing rate, β_s, increases by an order of magnitude when temperature is increased to close to the glass transition[4], so it seems possible that variations in healing rate—due to high pressures and temperatures or fault chemistry—could affect fault properties more profoundly than variations in recurrence time. If deeper sections of the fault are more healed than shallower fault sections, this might cause those parts to radiate more high-frequency energy when ruptured in a large earthquake.

The healing-related spectral changes observed in this study demonstrate how earthquake spectra are determined not simply by static fault strength or total fault slip, but by the manner in which slip occurs. Fault sections that heal rapidly or faults that heal for a long time, such as those associated with intraplate earthquakes in low-strain-rate environments, will produce higher-frequency earthquakes. In contrast, fault sections composed of materials that do not heal, such as smectite, a clay mineral found in the creeping section of the San Andreas fault[28] and in subduction zones[29], will slip slowly and smoothly.

METHODS SUMMARY

Sliding surfaces of laboratory specimens were milled flat and then roughened by hand lapping with either #60 grit or #600 grit abrasive, producing surface roughnesses referred to as rough and smooth, respectively. Shear force, F_S, was measured with a load cell located between the loading screw and the slider block. Fault slip, δ, was measured at both the leading edge (δ_L) and the trailing edge (δ_T) of the slider using eddy current sensors mounted on the samples. The loading screw was turned by hand; consequently, v_{LP} was not precisely controlled but was calculated from δ_T, F_S and the apparatus and specimen stiffness, which was constant for each run. Power spectral estimates were obtained by Fourier-transforming a 65.5-ms (laboratory) or 3.5–4-s (field) signal centred on the first arrival and tapered with a Blackman–Harris window. Noise spectra were obtained similarly. Only data with signal-to-noise ratio of at least 6 dB was used. LabEQs were recorded with a Panametrics V103 sensor located 80 mm from the laboratory fault. RESs near Parkfield were recorded as 250-Hz velocity seismograms by the borehole High-Resolution Seismic Network. Only vertical-component records were used for this study. RES detections and locations follow ref. 22. For each station and each RES, station averages were calculated by linearly averaging spectra from events cleanly recorded by all stations. We computed relative spectral amplitudes by dividing spectra of individual recordings by the station average. Relative spectral ratios were obtained from the ratio of relative spectral amplitudes at high frequencies (75–85 Hz) to those at lower frequencies (5–15 Hz). A different choice of high-frequency band (for example 65–75 Hz) does not affect the results.

Full Methods and any associated references are available in the online version of the paper.

Received 25 April; accepted 14 August 2012.

1. Dieterich, J. H. Time-dependent friction in rocks. *J. Geophys. Res.* **77,** 3690–3697 (1972).
2. Dieterich, J. H. & Kilgore, B. D. Direct observations of frictional contacts—new insights for state-dependent properties. *Pure Appl. Geophys.* **143,** 283–302 (1994).
3. Scholz, C. *The Mechanics of Earthquakes and Faulting* 81–100 (Cambridge Univ. Press, 2002).
4. Berthoud, P. & Baumberger, T. G'Sell, C. & Hiver, J.-M. Physical analysis of the state- and rate-dependent friction law: static friction. *Phys. Rev. B* **59,** 14313–14327 (1999).
5. Kiser, E. & Ishii, M. The 2010 Mw 8.8 Chile earthquake: Triggering on multiple segments and frequency-dependent rupture behavior. *Geophys. Res. Lett.* **38,** L07301 (2011).
6. Lay, T. *et al.* Depth-varying rupture properties of subduction zone megathrust faults. *J. Geophys. Res.* **117,** B04311 (2012).
7. Meng, L., Inbal, A. & Ampuero, J.-P. A window into the complexity of the dynamic rupture of the 2011 Mw 9 Tohoku-Oki earthquake. *Geophys. Res. Lett.* **38,** L00G07 (2011).
8. Karner, S. L. & Marone, C. in *Geocomplexity and the Physics of Earthquakes* (eds Rundle, J. B., Turcotte, D. and Klein, W.) 187–198 (Geophys. Monogr. Ser. 120, AGU, 2000).
9. Chen, T. & Lapusta, N. Scaling of small repeating earthquakes explained by interaction of seismic and aseismic slip in a rate and state fault model. *J. Geophys. Res.* **114,** B01311 (2009).
10. Chen, K. H., Bürgmann, R., Nadeau, R. M., Chen, T. & Lapusta, N. Postseismic variations in seismic moment and recurrence interval of repeating earthquakes. *Earth Planet. Sci. Lett.* **299,** 118–125 (2010).
11. Vidale, J. E., Ellsworth, W. L., Cole, A. & Marone, C. Variations in rupture process with recurrence interval in a repeated small earthquake. *Nature* **368,** 624–626 (1994).
12. Peng, Z. & Vidale, J. E. Marone, C. & Rubin, A. Systematic variations in recurrence interval and moment of repeating aftershocks. *Geophys. Res. Lett.* **32,** L15301 (2005).
13. Dieterich, J. H. & Conrad, G. Effect of humidity on time and velocity-dependent friction in rocks. *J. Geophys. Res.* **89,** 4196–4202 (1984).
14. Cox, S. F. & Paterson, M. S. Experimental dissolution precipitation creep in quartz aggregates at high temperatures. *Geophys. Res. Lett.* **18,** 1401–1404 (1991).
15. Li, Q., Tullis, T. E., Goldsby, D. & Carpick, R. W. Frictional ageing from interfacial bonding and the origins of rate and state friction. *Nature* **480,** 233–236 (2011).
16. Rabinowitz, E. *Friction and Wear of Materials* (Wiley, 1965).
17. Heslot, F., Baumberger, T., Perrin, B., Caroli, B. & Caroli, C. Creep, stick–slip, and dry friction dynamics: experiments and a heuristic model. *Phys. Rev. E* **49,** 4973–4988 (1994).
18. Rice, J. R. & Cocco, M. in *Tectonic Faults: Agents of Change on a Dynamic Earth* (eds Handy, M. R., Hirth, G. & Hovius, N.) 99–137 (MIT Press, 2007).
19. Wu, F. T., Thomson, K. C. & Kuenzler, H. Stick-slip propagation velocity and seismic source mechanism. *Bull. Seismol. Soc. Am.* **62,** 1621–1628 (1972).
20. McLaskey, G. C. & Glaser, S. D. Micromechanics of asperity rupture during laboratory stick slip experiments. *Geophys. Res. Lett.* **38,** L12302 (2011).
21. Beeler, N. M., Hickman, S. H. & Wong, T.-f. Earthquake stress drop and laboratory-inferred interseismic strength recovery. *J. Geophys. Res.* **106,** 30,701–30,713 (2001).
22. Nadeau, R. M. & McEvilly, T. V. Fault slip rates at depth from recurrence intervals of repeating microearthquakes. *Science* **285,** 718–721 (1999).
23. Nadeau, R. M., Michalini, A., Uhrhammer, R. A., Dolenc, D. & McEvilly, T. V. Detailed kinematics, structure and recurrence of micro-seismicity in the SAFOD target region. *Geophys. Res. Lett.* **31,** L12S08 (2004).
24. Rubinstein, J. L. & Beroza, G. C. Depth constraints on nonlinear strong ground motion. *Geophys. Res. Lett.* **32,** L14313 (2005).
25. Johnson, L. An earthquake model with interacting asperities. *Geophys. J. Int.* **182,** 1339–1373 (2010).
26. Dreger, D. Nadeau, R. M. & Chung, A. Repeating earthquake finite source models: strong asperities revealed on the San Andreas fault. *Geophys. Res. Lett.* **34,** L23302 (2007).
27. Page, M., Dunham, E. & Carlson, J. M. Distinguishing barriers and asperities in near-source ground motion. *J. Geophys. Res.* **110,** B11302 (2005).
28. Carpenter, B. M., Marone, C. & Saffer, D. M. Weakness of the San Andreas fault revealed by samples from the active fault zone. *Nature Geosci.* **4,** 251–254 (2011).
29. Saffer, D. M. & Marone, C. Comparison of smectite- and illite rich gouge frictional properties: application to the updip limit of the seismogenic zone along subduction megathrusts. *Earth Planet. Sci. Lett.* **215,** 219–235 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements This paper was improved by suggestions from R. Bürgmann and constructive reviews by T. Tullis, C. Marone, W. Ellsworth and N. Beeler. High-Resolution Seismic Network data was provided by the Berkeley Seismological Laboratory and NCEDC. Research was supported by the US NSF GRF and NSF grants CMMI-1131582, EAR-0738342 and EAR-0910322. This is BSL contribution #11-12.

Author Contributions G.C.M. and S.D.G. developed the laboratory experiments. R.M.N. developed and maintained repeating-earthquake catalogues. A.M.T. and G.C.M. performed analysis of the RESs at Parkfield. G.C.M. performed analysis of LabEQs and wrote the manuscript, with contributions from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.C.M. (gmclaskey@usgs.gov).

LETTER

doi:10.1038/nature11585

Combined pesticide exposure severely affects individual- and colony-level traits in bees

Richard J. Gill[1], Oscar Ramos-Rodriguez[1] & Nigel E. Raine[1]

Reported widespread declines of wild and managed insect pollinators have serious consequences for global ecosystem services and agricultural production[1–3]. Bees contribute approximately 80% of insect pollination, so it is important to understand and mitigate the causes of current declines in bee populations[4–6]. Recent studies have implicated the role of pesticides in these declines, as exposure to these chemicals has been associated with changes in bee behaviour[7–11] and reductions in colony queen production[12]. However, the key link between changes in individual behaviour and the consequent impact at the colony level has not been shown. Social bee colonies depend on the collective performance of many individual workers. Thus, although field-level pesticide concentrations can have subtle or sublethal effects at the individual level[8], it is not known whether bee societies can buffer such effects or whether it results in a severe cumulative effect at the colony level. Furthermore, widespread agricultural intensification means that bees are exposed to numerous pesticides when foraging[13–15], yet the possible combinatorial effects of pesticide exposure have rarely been investigated[16,17]. Here we show that chronic exposure of bumblebees to two pesticides (neonicotinoid and pyrethroid) at concentrations that could approximate field-level exposure impairs natural foraging behaviour and increases worker mortality leading to significant reductions in brood development and colony success. We found that worker foraging performance, particularly pollen collecting efficiency, was significantly reduced with observed knock-on effects for forager recruitment, worker losses and overall worker productivity. Moreover, we provide evidence that combinatorial exposure to pesticides increases the propensity of colonies to fail.

The majority of studies to date have focused on pesticide exposure in honeybees, but bumblebees are also crucial pollinators and have smaller colonies, making them ideally suited to investigate effects at both the individual (worker) and colony level. This study mimicked a realistic scenario in which 40 early-stage bumblebee (*Bombus terrestris*) colonies received long-term (4-week) exposure to two widely used pesticides frequently encountered when foraging on flowering crops, the neonicotinoid imidacloprid and the pyrethroid λ-cyhalothrin. Imidacloprid is a systemic pesticide found in all plant tissues, including the pollen and nectar consumed by bees (oral exposure[18–20]). λ-cyhalothrin is sprayed directly on to crops, including their flowers, to which bees will be topically exposed (details in Supplementary Information). Foraging bees are thus simultaneously exposed to both chemicals in the field, making them excellent candidates to investigate the potential for combinatorial effects of pesticide exposure. Using a split block design (see Methods), we monitored colonies exposed to each pesticide independently and in combination (ten control colonies, ten exposed to imidacloprid (I), ten exposed to λ-cyhalothrin (LC) and ten exposed to I and LC (mix = M)). Imidacloprid (dissolved in 40% sucrose solution) was provided at a concentration (10 p.p.b. (parts per billion (10^9)) within the range found in crop nectar and pollen in the field[9,21]. λ-cyhalothrin was administered following label guidance for field-spray application (see Supplementary Information). Bees were able to forage in the field, providing a realistic and demanding behavioural setting, and the foraging behaviour of individual workers was recorded using radio frequency identification (RFID) tagging technology[10,11,22] (Supplementary Figs 1 and 2). Colonies were motivated to forage because we provided them with no pollen and limited amounts of sucrose solution.

During colony development, the production of workers (and their survival) is vital to colony success because workers provide the labour (for example, brood care and foraging) for the colony. Total worker production at the end of the experiment was significantly lower in imidacloprid-treated colonies (reduced by 27% in I and 9% in M colonies) compared to control colonies (mean (± s.e.m.) workers per colony, I = 19.7 ± 3.0, M = 24.4 ± 3.2 versus control = 27.0 ± 4.0; linear mixed effects model (LMER), I, $Z = -3.71$, $P < 0.001$; M, $Z = -2.62$, $P = 0.009$; Fig. 1a). Two of the forty colonies, both M colonies, did not survive the experiment (they 'failed' after 3 and 8 days; see Supplementary Information), a colony failure rate significantly higher than other treatments (Fisher's Exact test: mid-P correction = 0.029). These two colonies were excluded from statistical analyses to provide a conservative assessment of worker production in M colonies (when included in analysis = 20.0 ± 3.9 workers). During the experiment, 223 (21% of total) workers were found dead inside nest boxes. On average, 36 ± 7.3% and 39 ± 7.5% of workers from LC and M colonies, respectively, died in the nest box; a figure four times higher than control (9 ± 3.4%) colonies (LMER, LC, $t = 4.31$, $P < 0.001$; M, $t = 4.23$, $P < 0.001$; Fig. 1b). Moreover, 43% of the workers found dead in LC and M colonies lived fewer than 4 days after eclosion—an apparent waste of resources required for future colony growth given that such young members are unlikely to have contributed any work (for example, foraging) to offset the resources invested to produce them. Queen loss occurred in 14 colonies, although loss rate did not differ significantly among treatments (control = 4; I = 5; LC = 2; M = 3; Fisher's exact test: mid-P-correction = 0.40) and we accounted for queen loss in our analyses (see Supplementary Information).

Daily counts of newly eclosed bees showed that worker production in I colonies did not become significantly lower than control colonies until the end of week 2, and for M colonies until the end of week 4 (Fig. 1c; see Supplementary Information and Supplementary Table 1). Daily counts of dead bees also revealed that worker mortality in LC colonies did not become significantly higher than that in control colonies until the end of week 3, but worker mortality in M colonies became significantly higher than that in control colonies as early as the end of week 1. The delayed effect of imidacloprid exposure on worker productivity in I and M colonies coincides with the time taken by workers to develop from egg to adult (approximately 22 days), suggesting that the observed effect is a result of imidacloprid on brood development. Indeed, the total number of larvae and pupae combined that were found in colonies at the end of the experiment ('brood number') was significantly lower in I and M colonies compared to control colonies (LMER, I, $Z = -6.23$, $P < 0.001$; M, $Z = -5.60$, $P < 0.001$). Overall, this represented a 22% reduction in brood production in I colonies and a 7% reduction in M colonies (mean (± s.e.m.) brood number, I = 36 ± 8.0, M = 43 ± 11.7 (including failed colonies:

---

[1] School of Biological Sciences, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK.

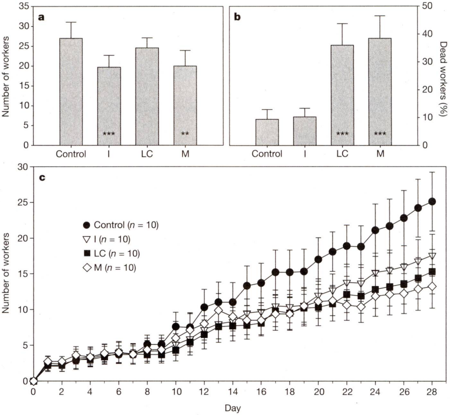


Figure 1 | Worker production and mortality. **a**, Mean (± s.e.m.) number of workers per colony that eclosed by the end of the experiment. **b**, Mean percentage of workers per colony found dead inside the nest box by the end of the experiment. **c**, Colony growth shown by daily counts of the cumulative number of workers eclosed minus the cumulative number of workers found dead (mean (± s.e.m.) per colony). Data shown on the x axis indicate the number of days since the start of the experiment (day 1 = 24 h after the start of experimentation). M treatment includes the two collapsed colonies. $^{*}P \leq 0.05$, $^{**}P \leq 0.01$, $^{***}P \leq 0.001$ (comparison with control).

M = 39 ± 9.6) versus control = 46 ± 9.7). Despite this, there was no significant difference in the mass of the wax nest structure (see Supplementary Information for details) across treatments at the end of the experiment (LMER, I, $t = -1.12$, $P = 0.27$; M, $t = -1.22$, $P = 0.23$; Supplementary Fig. 3) indicating that I and M colonies attempted to raise similar brood numbers but that a lower proportion of larvae and pupae survived to eclosion.

Although imidacloprid could be directly affecting brood (physiological) development, it could also indirectly affect the brood by causing changes to colony behaviour and/or structure: for example, changes to foraging behaviour leading to food limitation[23,24]. We tested this hypothesis by studying worker foraging performance using RFID technology to automatically record the exact time workers left or entered each colony (Supplementary Figs 1 and 2). Overall, we collected data from 259 recognized foragers from 32 colonies (n colonies: control = 7; I = 10; LC = 8; M = 7) making 8,751 foraging bouts (median (interquartile range) per worker = 23 (10–44); for criteria used to classify foragers and foraging bouts see Methods). We examined whether pesticide treatment affected foraging activity and forager recruitment. We found that foragers from M colonies performed fewer foraging bouts compared to control colonies (LMER, $t = -2.55$, $P = 0.011$; Fig. 2a), and that there were significantly more foragers in both I and M colonies compared to control colonies over the 4 weeks (LMER, I, $Z = 4.20$, $P < 0.001$; M, $Z = 3.49$, $P < 0.001$; Fig. 2a). The higher number of foragers in I and M colonies (compared to control) is unlikely to be due to either pesticide causing a significant repellent or anti-feedant effect (this corroborates the lack of published evidence for pyrethroid repellency in bumblebees despite reports of pyrethroids being repellent to honeybees[25]). This is because workers did not have to visit the feeder, as they could forage for nectar outside, yet we found no difference among treatments in the amount of sucrose collected from feeders (LMER, $t \leq 1.63$, $P \geq 0.11$; Supplementary Fig. 6).

Given that I and M colonies recruited higher numbers of workers to forage compared to control colonies, we evaluated whether this was a response to reduced individual foraging efficiency by monitoring pollen foraging performance and observing the size of pollen loads (load size scored as: small = 1, medium = 2, large = 3; see Methods) brought back by foragers (n = 20 h of observation per colony). Crucially, imidacloprid-exposed foragers returned with significantly smaller pollen loads per foraging bout compared to control colonies (LMER, I, $t = -3.31$, $P = 0.0011$; M, $t = -3.38$, $P < 0.001$; Fig. 2b). Imidacloprid-exposed foragers collected pollen successfully in a significantly lower percentage of their foraging bouts (mean (± s.e.m.), I = 59 ± 7.3%, M = 55 ± 8.6% versus control = 82 ± 5.8%; LMER, I, $t = -3.16$, $P = 0.0018$; M, $t = -3.05$, $P = 0.0026$; Supplementary Fig. 4) and we also found that the average duration of successful foraging bouts (during which pollen was collected) was significantly longer for imidacloprid-exposed foragers than for control foragers (LMER, I, $t = 2.10$, $P = 0.037$; M, $t = 2.87$, $P = 0.005$; Fig. 2c). Together, these data show that imidacloprid-exposed workers were significantly less efficient at collecting pollen in the field.

A consequence of recruiting a greater number of workers to forage is that it increases the proportion of colony workforce going outside to undertake a potentially hazardous task[22]. Indeed, our RFID data show the number of foragers per colony was significantly correlated with the number of workers leaving the colony and getting 'lost' outside (that is, workers that did not return: Spearman's Rank, $\rho = 0.801$, $P < 0.001$; Supplementary Fig. 5). Consequently, we found that on average the percentage of workers getting lost in I and M colonies was 50% and 55% higher than control colonies (I = 30 ± 3.1%, M = 31 ± 5.3% versus control = 20 ± 2.9%; LMER, I, $t = 2.83$, $P = 0.008$; M, $t = 2.26$,

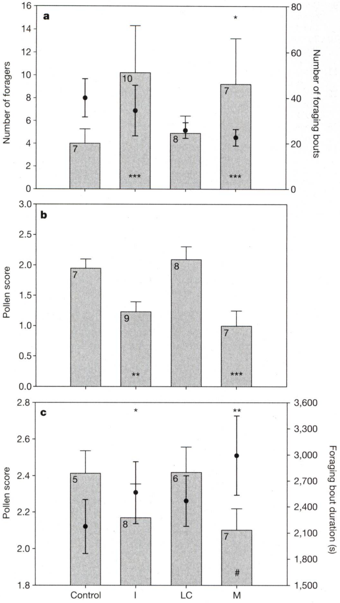


Figure 2 | Foraging performance. **a**, Mean (± s.e.m.) number of foragers per colony (column), and foraging bouts per worker per colony (filled circles: $n = 259$ foragers). **b**, Mean pollen score per worker per colony for all observed foraging bouts ($n = 228$ foragers). **c**, Mean pollen score per successful (pollen) foraging bout for each worker per colony (column), and mean duration of successful foraging bouts per worker per colony (filled-circles) ($n = 147$ foragers). n colonies shown in top left corner of columns. Significant differences from control treatment for column data are shown at the bases of columns, and for filled-circle data are shown above columns (**a** and **c**). #$P \leq 0.1$, *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$ (comparison with control).

$P = 0.03$). Furthermore, when considering worker mortality and losses combined over the 4 weeks (mean (± s.e.m.): I = 41 ± 4.2%, LC = 51 ± 6.8%, M = 69 ± 7.1% versus control = 30 ± 5.0%, LMER, I, $t = 1.79$, $P = 0.08$; LC, $t = 3.25$, $P = 0.0026$; M, $t = 5.24$, $P < 0.001$; Table 1 and Fig. 3), we found that colonies treated with both pesticides (M) suffered most severely. Moreover, M colonies had significantly higher overall worker losses than either I colonies (LMER, $t = -3.69$, $P < 0.001$) or LC colonies (LMER, $t = -2.31$, $P = 0.027$).

We have shown that imidacloprid exposure at concentrations that can be found in the pollen and nectar of flowering crops causes impairment to pollen foraging efficiency, leading to increased colony demand for food as shown by increased worker recruitment to forage. However, imidacloprid-treated colonies (I and M) were still unable to collect as much pollen as control colonies. Such pollen constraints, coupled with a higher number of workers undertaking foraging rather than brood care, seemed to affect brood development, resulting in reduced worker production that can only exacerbate the problem of having an impaired colony workforce. These findings show a mechanistic explanation to link recently reported effects on individual worker behaviour[10,11,26–29] and colony queen production[12] as a result of neonicotinoid exposure. Moreover, exposure to a second pesticide λ-cyhalothrin (pyrethroid) applied at label-guideline concentration for crop use caused additional worker mortality in this study highlighting another potential risk. Bee colonies typically encounter several classes of pesticides when foraging in the field[13–15], potentially exposing them to a range of combinatorial effects. Indeed, M colonies in our study were consistently negatively affected in all our measures of worker behaviour, suffered the highest overall worker losses (worker mortality and forager losses), which were twice as great as for control colonies, and two colonies did in fact fail (Table 1).

Pesticide-label-guidance concentrations and application rates are approved on the basis of ecotoxicological tests using single pesticides and set at a level for field use deemed 'sublethal' (below a dose lethal to 50% of animals tested (LD_{50})). However, the risk of exposure to multiple pesticides, or of the same pesticide being applied to different (adjacent) crops, is currently not considered when evaluating the safety of pesticides for bees. Given the serious impacts on M colonies it is concerning that pesticide products containing mixtures of neonicotinoids and pyrethroids are in current use[18]. At present there are also no guidelines for testing chronic or sublethal effects of pesticides on

Table 1 | Summary of observed pesticide effects for each treatment group (I, LC or M) in comparison to the control group

Effect level	Effect type	I	LC	M
Effects on individual behaviour	Number of foragers	+	ND	+
	Foraging bout frequency	ND	ND	−
	Amount of pollen collected	−	ND	−
	Duration of pollen foraging bouts	+	ND	+
Effects at colony level	Worker production	−	ND	−
	Brood number	−	ND	−
	Nest structure mass	ND	ND	ND
	Worker mortality	ND	+	+
	Worker loss	+	−	+
	Worker mortality & loss	ND	+	+
	Colony failure (*n* failed/*n* survived)	0/10	0/10	2/8

Significant decrease (−), significant increase (+) and no detected effect (ND) at the 5% significance level.

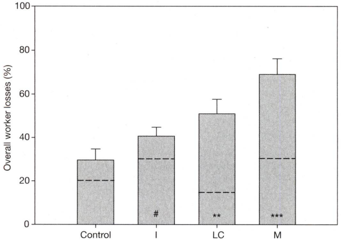


Figure 3 | Overall worker losses. Mean (± s.e.m.) overall percentage of workers lost per colony, including workers lost outside (below the dashed line) and worker mortality (dead workers found in nest box; above the dashed line), during the 4-week experiment. $n = 40$ colonies. #$P \leq 0.1$, **$P \leq 0.01$, ***$P \leq 0.001$ (comparison with control).

bees[30], and considering that we did not detect significant effects until 2 to 4 weeks into our study, the current European and Mediterranean Plant Protection Organisation (EPPO) and Organization for Economic Co-operation and Development (OECD) guideline of a maximum exposure of 96 h (for testing acute effects of pesticides on honeybees) appears to be insufficient. Our results emphasize the importance of recent recommendations by the European Food Safety Authority (EFSA) Panel on Plant Protection Products and their Residues (http://www.efsa.europa.eu/en/efsajournal/pub/2668.htm) proposing the need for longer term toxicity testing on both adult bees and larvae, new protocols to detect cumulative toxicity effects and separate risk assessment schemes for different bee species. Our findings have clear implications for the conservation of insect pollinators in areas of agricultural intensification, particularly social bees with their complex social organization and dependence on a critical threshold of workers performing efficiently to ensure colony success.

METHODS SUMMARY

Each colony contained a queen and ten or fewer workers at the start of the experiment, with no significant difference among treatments in worker number (Kruskal–Wallis: $H = 0.26$, $P = 0.97$). Colonies were housed in two-chambered nest boxes, with the rear chamber housing the nest and front chamber used for pesticide exposure (Supplementary Figs 1 and 6). Nest boxes were kept in the laboratory but connected via an outlet tube to the outside to allow natural foraging. Foraging activity of tagged workers was automatically recorded by RFID readers placed at the entrance to each nest box (Supplementary Fig. 2). The food chamber housed a feeder containing a specified volume (averaging 13ml) of control sucrose solution (control and LC) or 10 p.p.b. imidacloprid sucrose solution (I and M) provided every 2 to 3 days (Supplementary Table 2). The feeder was placed in a Petri dish lined with filter paper that was sprayed once at the start of each week with 0.69 ± 0.046 ml of control solution (control and I) or 37.5 p.p.m. (parts per million (10^6)) λ-cyhalothrin solution (LC and M). Workers walking across the filter paper to the feeder had contact exposure to λ-cyhalothrin (LC and M), and oral exposure to imidacloprid (I and M) when feeding. Colonies were not provided with pollen to motivate foraging behaviour. All workers were RFID tagged, with new workers tagged within 3 days of eclosion (Supplementary Fig. 2). We classified a foraging bout as a period of at least 5 minutes between a worker leaving and returning to a colony, and a forager as a worker that performed at least 4 foraging bouts. Pollen foraging was observed for 1 hour per colony per day (5 days per week) recording the presence and size of pollen loads collected (Supplementary Table 2). Colonies were frozen at the end of the experiment; the number of workers (and tag identifications) and brood was counted, and the mass of the nest structure was recorded.

Full Methods and any associated references are available in the online version of the paper.

Received 19 April; accepted 13 September 2012.

Published online 21 October 2012.

1. Biesmeijer, J. C. *et al.* Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science* **313,** 351–354 (2006).
2. Klein, A. M. *et al.* Importance of pollinators in changing landscapes for world crops. *Proc. R. Soc. B* **274,** 303–313 (2007).
3. Kremen, C. *et al.* Pollination and other ecosystem services produced by mobile organisms: a conceptual framework for the effects of land-use change. *Ecol. Lett.* **10,** 299–314 (2007).
4. Potts, S. G. *et al.* Global pollinator declines: trends, impacts and drivers. *Trends Ecol. Evol.* **25,** 345–353 (2010).
5. Oldroyd, B. P. What's killing American honey bees? *PLoS Biol.* **5,** e168 (2007).
6. Brown, M. J. F. & Paxton, R. J. The conservation of bees: a global perspective. *Apidologie* **40,** 410–416 (2009).
7. Thompson, H. M. Behavioural effects of pesticides in bees - their potential for use in risk assessment. *Ecotoxicology* **12,** 317–330 (2003).
8. Desneux, N., Decourtye, A. & Delpuech, J. M. The sublethal effects of pesticides on beneficial arthropods. *Annu. Rev. Entomol.* **52,** 81–106 (2007).
9. Cresswell, J. E. A meta-analysis of experiments testing the effects of a neonicotinoid insecticide (Imidacloprid) on honey bees. *Ecotoxicology* **20,** 149–157 (2011).
10. Schneider, C. W., Tautz, J., Grünewald, B. & Fuchs, S. RFID tracking of sublethal effects of two neonicotinoid insecticides on the foraging behavior of *Apis mellifera*. *PLoS ONE* **7,** e30023 (2012).
11. Henry, M. *et al.* A common pesticide decreases foraging success and survival in honey bees. *Science* **336,** 348–350 (2012).
12. Whitehorn, P. R., O'Connor, S., Wackers, F. L. & Goulson, D. Neonicotinoid pesticide reduces bumble bee colony growth and queen production. *Science* **336,** 351–352 (2012).
13. Johnson, R. M., Ellis, M. D., Mullin, C. A. & Frazier, M. Pesticides and honey bee toxicity — USA. *Apidologie* **41,** 312–331 (2010).
14. Mullin, C. A. *et al.* High levels of miticides and agrochemicals in North American apiaries: implications for honey bee health. *PLoS ONE* **5,** e9754 (2010).
15. Krupke, C. H., Hunt, G. J., Eitzer, B. D., Andino, G. & Given, K. Multiple routes of pesticide exposure for honey bees living near agricultural fields. *PLoS ONE* **7,** e29268 (2012).
16. Johnson, R. M., Pollock, H. S. & Berenbaum, M. R. Synergistic interactions between in-hive miticides in *Apis mellifera*. *J. Econ. Entomol.* **102,** 474–479 (2009).
17. Pilling, E. D. & Jepson, P. C. Synergism between EBI fungicides and a pyrethroid insecticide in the honeybee (*Apis mellifera*). *Pestic. Sci.* **39,** 293–297 (1993).
18. Elbert, A., Haas, M., Springer, B., Thielert, W. & Nauen, R. Applied aspects of neonicotinoid uses in crop protection. *Pest Manag. Sci.* **64,** 1099–1105 (2008).
19. Rortais, A., Arnold, G., Halm, M. P. & Touffet-Briens, F. Modes of honeybees exposure to systemic insecticides: estimated amounts of contaminated pollen and nectar consumed by different categories of bees. *Apidologie* **36,** 71–83 (2005).
20. Chauzat, M. P. *et al.* A survey of pesticide residues in pollen loads collected by honey bees in France. *J. Econ. Entomol.* **99,** 253–262 (2006).
21. Blacquière, T., Smagghe, G., van Gestel, C. A. M. & Mommaerts, V. Neonicotinoids in bees: a review on concentrations, side-effects and risk assessment. *Ecotoxicology* **21,** 973–992 (2012).
22. Molet, M., Chittka, L., Stelzer, R. J., Streit, S. & Raine, N. E. Colony nutritional status modulates worker responses to foraging recruitment pheromone in the bumblebee *Bombus terrestris*. *Behav. Ecol. Sociobiol.* **62,** 1919–1926 (2008).
23. Schmid-Hempel, R. & Schmid-Hempel, P. Colony performance and immunocompetence of a social insect, *Bombus terrestris*, in poor and variable environments. *Funct. Ecol.* **12,** 22–30 (1998).
24. Pelletier, L. & McNeil, J. N. The effect of food supplementation on reproductive success in bumblebee field colonies. *Oikos* **103,** 688–694 (2003).
25. Thompson, H. & Wilkins, S. Assessment of the synergy and repellency of pyrethroid/fungicide mixtures. *Bull. Insectol.* **56,** 131–134 (2003).
26. Bortolotti, L. *et al.* Effects of sub-lethal imidacloprid doses on the homing rate and foraging activity of honey bees. *Bull. Insectol.* **56,** 63–67 (2003).
27. Decourtye, A., Devillers, J., Cluzeau, S., Charreton, M. & Pham-Delègue, M. H. Effects of imidacloprid and deltamethrin on associative learning in honeybees under semi-field and laboratory conditions. *Ecotoxicol. Environ. Saf.* **57,** 410–419 (2004).
28. Mommaerts, V. *et al.* Risk assessment for side-effects of neonicotinoids against bumblebees with and without impairing foraging behavior. *Ecotoxicology* **19,** 207–215 (2010).
29. Yang, E. C., Chuang, Y. C., Chen, Y. L. & Chang, L. H. Abnormal foraging behavior induced by sublethal dosage of Imidacloprid in the Honey bee (Hymenoptera: Apidae). *J. Econ. Entomol.* **101,** 1743–1748 (2008).
30. Halm, M. P., Rortais, A., Arnold, G., Taséi, J. N. & Rault, S. New risk assessment approach for systemic insecticides: the case of honey bees and imidacloprid (Gaucho). *Environ. Sci. Technol.* **40,** 2448–2454 (2006).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M.J.F. Brown, M. Clook, J. Culverhouse, A. Dixon, M. Fürst, D. Garthwaite, A. Horsell, V. Jansen and I. Pedroso-Rovira for comments and technical assistance, and Syngenta Bioline Bees for supplying colonies. The study was supported by the Insect Pollinator Initiative (funded under the auspices of the Living with Environmental Change programme, Biotechnology and Biological Sciences Research Council (BBSRC), Wellcome Trust, Scottish Government, Department for Environment, Food and Rural Affairs (DEFRA) and Natural Environment Research Council (NERC): grant BB/I000178/1).

Author Contributions R.J.G., O.R.-R. and N.E.R. carried out the experiment; R.J.G. and N.E.R. designed the experiment and wrote the paper; N.E.R. conceived the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.J.G. (richard.gill@rhul.ac.uk) or N.E.R. (nigel.raine@rhul.ac.uk).

LETTER

doi:10.1038/nature11523

Intrinsically determined cell death of developing cortical interneurons

Derek G. Southwell[1,2,3]†, Mercedes F. Paredes[2,4], Rui P. Galvao[2]†, Daniel L. Jones[1,2], Robert C. Froemke[5]†, Joy Y. Sebe[2], Clara Alfaro-Cervello[6]†, Yunshuo Tang[2,3,7], Jose M. Garcia-Verdugo[6], John L. Rubenstein[8], Scott C. Baraban[1,2] & Arturo Alvarez-Buylla[1,2]

Cortical inhibitory circuits are formed by γ-aminobutyric acid (GABA)-secreting interneurons, a cell population that originates far from the cerebral cortex in the embryonic ventral forebrain. Given their distant developmental origins, it is intriguing how the number of cortical interneurons is ultimately determined. One possibility, suggested by the neurotrophic hypothesis[1–5], is that cortical interneurons are overproduced, and then after their migration into cortex the excess interneurons are eliminated through a competition for extrinsically derived trophic signals. Here we characterize the developmental cell death of mouse cortical interneurons *in vivo*, *in vitro* and after transplantation. We found that 40% of developing cortical interneurons were eliminated through Bax (Bcl-2-associated X)-dependent apoptosis during postnatal life. When cultured *in vitro* or transplanted into the cortex, interneuron precursors died at a cellular age similar to that at which endogenous interneurons died during normal development. Over transplant sizes that varied 200-fold, a constant fraction of the transplanted population underwent cell death. The death of transplanted neurons was not affected by the cell-autonomous disruption of TrkB (tropomyosin kinase receptor B), the main neurotrophin receptor expressed by neurons of the central nervous system[6–8]. Transplantation expanded the cortical interneuron population by up to 35%, but the frequency of inhibitory synaptic events did not scale with the number of transplanted interneurons. Taken together, our findings indicate that interneuron cell death is determined intrinsically, either cell-autonomously or through a population-autonomous competition for survival signals derived from other interneurons.

We first characterized the developmental cell death of cortical interneurons by measuring the expression of the apoptotic marker, cleaved caspase-3, in mice in which neurons expressing GAD67 were labelled with green fluorescent protein (GFP; GAD67–GFP mice)[9] (Fig. 1a). The number of cleaved caspase-3-labelled neocortical GAD67–GFP neurons increased from postnatal days 1 to 5 (P1 to P5), reached a maximum at about P7, and declined towards zero by about P15 (Fig. 1b; analysis of variance (ANOVA), $F = 84.0$ and $P < 0.0001$). Most (75%) cleaved caspase-3-positive cells were observed between P7 and P11 (Fig. 1b), about 11–18 days after the cells were produced in the embryonic ventral forebrain[10]. The temporal profile of expression of cleaved caspase-3 in GAD67–GFP cells was similar to that observed across the total cellular population of the neocortex (Supplementary Fig. 2), which may preserve the relative sizes of different cellular populations[11]. Because the GAD67–GFP knock-in decreases brain GABA content by about 20–40% (ref. 9), we examined whether this in turn affected cell death in GAD67–GFP mice. Across the entire cellular population of the neocortex, neither the temporal profile nor the extent of apoptosis was significantly different between GAD67–GFP mice and wild-type mice (Supplementary Fig. 3).

We next measured the GAD67–GFP population size during postnatal life and adulthood (Fig. 1c). The number of GAD67–GFP neurons reached a maximum at about P5 (mean ± s.e.m. $(1.65 \pm 0.03) \times 10^6$ cells), and then declined by about 40% during the period of interneuron cell death (Fig. 1b), reaching a stable size of $(1.01 \pm 0.02) \times 10^6$ cells by P120 (mean; ANOVA, $F = 32.1$ and $P < 0.0001$). The developmental cell death of cortical interneurons depended on Bax function: at P7, when GAD67–GFP cell death reached a maximum in wild-type mice (Fig. 1b), GAD67–GFP cell death was nearly absent in *Bax*$^{-/-}$;GAD67–GFP mutants[12] (Fig. 1d; Student's *t*-test, $P = 0.0034$). Between P5 and P120 the cortical GAD67–GFP population did not decline in *Bax* mutants (Fig. 1e; ANOVA, $F = 2.28$, $P = 0.18$); at P120 the cortical interneuron population was 33% smaller in wild-type GAD67–GFP mice than in *Bax*$^{-/-}$;GAD67–GFP littermates ($(1.02 \pm 0.04) \times 10^6$ cells versus $(1.52 \pm 0.08) \times 10^6$ cells, respectively; Student's *t*-test, $P = 0.0041$). In wild-type and *Bax* mutant mice, similar proportions of GAD67–GFP neurons were labelled by parvalbumin, somatostatin, neuropeptide Y and calretinin (Supplementary Fig. 4), indicating that Bax-dependent cell death occurred uniformly across neurochemically defined interneuron subtypes. These findings indicate that *Bax*-dependent programmed cell death eliminates roughly 40% of neocortical interneurons during postnatal life.

After characterizing neocortical interneuron cell death *in vivo*, we examined whether neocortical interneurons undergo a similar pattern of cell death *in vitro*. We placed interneuron precursors from the embryonic day 13.5 (E13.5) GAD67–GFP medial ganglionic eminence (MGE) onto P0 to P2 neocortical feeder layers[13] (Fig. 2a) and quantified the expression of cleaved caspase-3 at various time points (Fig. 2b). GAD67–GFP neurons underwent cell death *in vitro*, with expression of cleaved caspase-3 reaching a maximum at 13 days (Fig. 2c; ANOVA, $F = 9.12$ and $P < 0.0001$). Roughly 66% of cell death occurred between 11 and 15 days *in vitro* (DIV), at about which time the GAD67–GFP cell number declined by about 30% (Fig. 2d; ANOVA, $F = 4.53$ and $P = 0.0012$). As previously mentioned, *in vivo* most interneuron cell death occurred between P7 and P11, when the developing cells were similarly between 11 and 18 days old (Fig. 1b). Interneuron cell death thus manifests *in vitro*, with a temporal pattern resembling that observed *in vivo*.

---

[1]Neuroscience Graduate Program, University of California, San Francisco, California 94143, USA. [2]Departments of Neuroscience and Neurosurgery, and the Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research; University of California, San Francisco, California 94143, USA. [3]Medical Scientist Training Program, University of California, San Francisco, California 94143, USA. [4]Department of Neurology, University of California, San Francisco, California 94143, USA. [5]Department of Otolaryngology, Coleman Memorial Laboratory and W.M. Keck Foundation Center for Integrative Neuroscience, University of California, San Francisco, California 94143, USA. [6]Instituto Cavanilles, Universidad de Valencia, CIBERNED, Valencia 46071, Spain. [7]Biomedical Sciences Graduate Program, University of California, San Francisco, California 94143, USA. [8]Department of Psychiatry and the Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California, San Francisco, California 94143, USA. †Present addresses: Department of Neurosurgery, Stanford University School of Medicine, Stanford, California 94305, USA (D.G.S.); Department of Molecular Biology, University of Oregon, Eugene, Oregon 97403, USA (R.P.G.); Molecular Neurobiology Program, The Helen and Martin Kimmel Center for Biology and Medicine at the Skirball Institute of Biomolecular Medicine, Departments of Otolaryngology, Physiology and Neuroscience, New York University School of Medicine, New York, New York 10016, USA (R.C.F.); Cambridge Centre for Brain Repair, Department of Clinical Neurosciences and Stem Cell Institute, University of Cambridge, Cambridge CB2 0PY, UK (C.A.-C.).

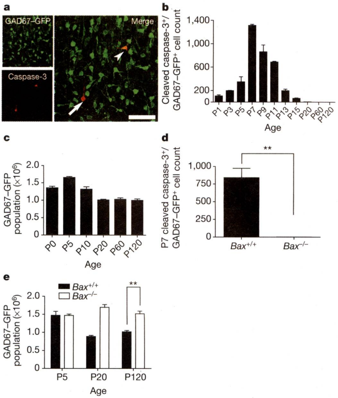


Figure 1 | Bax-dependent programmed cell death eliminates 40% of developing interneurons from the postnatal mouse neocortex. **a**, Expression of cleaved caspase-3 (red) observed in GAD67–GFP neurons (green; arrowhead) and other cell types (arrow) of the P7 neocortex. Scale bar, 100 μm (left) and 50 μm (right). **b**, Temporal profile of expression of cleaved caspase-3 in the neocortex of GAD67–GFP mice. Expression of cleaved caspase-3 is highest at P7, and declines to nearly undetectable levels by P15 (ANOVA, $F = 84.00$ and $P < 0.0001$; $n = 3$ per time point). **c**, Temporal profile of the neocortical GAD67–GFP population size. Between P5 and P20, the neocortical GAD67–GFP population decreases by about 40% (ANOVA, $F = 32.10$ and $P < 0.0001$; $n = 5$ per time point). **d**, The *Bax* mutation disrupts the developmental cell death of cortical interneurons. $Bax^{-/-}$ mice show a 99.8% decrease in the number of cells double labelled by cleaved caspase-3 and GAD67–GFP, compared with $Bax^{+/+}$;GAD67–GFP littermates (Student's *t*-test; $^{**}P < 0.01$; $n = 3$ per genotype). **e**, The neocortical GAD67–GFP population does not decrease in $Bax^{-/-}$ mice (ANOVA, $F = 2.28$ and $P = 0.18$). At P120, the neocortical GAD67–GFP population is roughly 33% smaller in wild-type mice (Student's *t*-test, $^{**}P < 0.01$; $n = 3$ per genotype at each time point). All error bars represent s.e.m.

We next transplanted embryonic interneuron precursors into the postnatal neocortex during the period of endogenous interneuron cell death[14,15]. We postulated that, if the timing of interneuron cell death reflected the maturation of interneurons into a trophic signal-dependent state, transplanted interneuron precursors would undergo developmental cell death asynchronously from endogenous interneurons. We transplanted 5×10^5 cells from the MGE of E13.5 to E14.5 β-actin:GFP mice[16] into P3 wild-type recipients (Supplementary Fig. 5) and then quantified the expression of cleaved caspase-3 at various time points after transplantation. Given that mouse gestation ends t about E19, the transplanted interneuron precursors were roughly 6–10 days younger than their endogenous counterparts[10]. As described previously[14,15,17], transplanted MGE cells dispersed in the recipient cortex, developed the morphological features of GABA-secreting interneurons (Supplementary Fig. 5) and formed synaptic contacts with recipient neurons (Supplementary Fig. 6). We did not observe labelling of the transplanted interneuron precursors by antigen Ki-67, indicating that the cells did not proliferate in the recipient (Supplementary Fig. 7). Expression of cleaved caspase-3 increased 200% in the transplanted population between 7 and 15 days after transplantation (DAT), reached a maximum at 15 DAT, then declined to undetectable levels by 45 DAT (Fig. 2e, f; ANOVA, $F = 17.79$ and $P < 0.0001$). By contrast, in endogenous cells of the recipient neocortex, expression of cleaved caspase-3 reached a relative maximum at 7 DAT, then declined roughly 80% between 7 and 15 DAT (Supplementary Fig. 8; ANOVA, $F = 401.20$ and $P < 0.0001$). The addition of transplanted cells did not affect endogenous cell death, because expression of cleaved caspase-3 was similar between hemispheres that received transplanted cells and hemispheres that received injections of vehicle medium (Supplementary Fig. 8; Student's *t*-test, $P = 0.76$ (7 DAT), $P = 0.83$ (15 DAT), $P = 0.89$ (25 DAT), $P = 0.67$ (45 DAT)). Transplanted interneuron cell death thus reached a maximum at about 15 DAT, when the transplanted cells reached a cellular age equivalent to that of endogenous interneurons

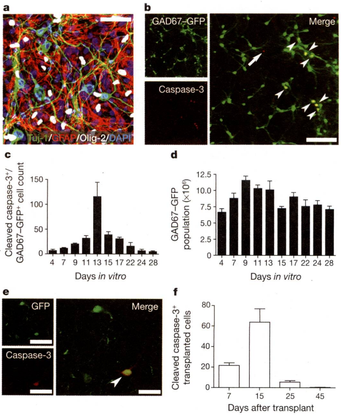


Figure 2 | *In vitro*, and after heterochronic transplantation, interneuron precursors undergo programmed cell death during a period defined by their intrinsic cellular age. **a**, Primary feeder layers prepared from P0 to P2 neocortex. At 14 DIV, the feeder layer contains neurons (Tuj-1, green), astrocytes (glial fibrillary acidic protein (GFAP), red) and oligodendrocytes (Olig-2, white). All cells are labelled by 4′,6-diamidino-2-phenylindole (DAPI, blue). Scale bar, 50 μm. **b**, At 14 DIV, double-labelled cells expressing cleaved caspase-3 (red) and GAD67–GFP (green; arrowheads) are observed along with cells singly labelled by cleaved caspase-3 (arrow). Scale bar, 200 μm. **c**, Temporal profile of expression of cleaved caspase-3 in GAD67–GFP neuronal cultures. Expression of cleaved caspase-3 is highest at 13 DIV (ANOVA, $F = 9.12$ and $P < 0.0001$). **d**, Temporal profile of the GAD67–GFP population size *in vitro*. The GAD67–GFP population increases in number between 4 and 9 DIV, probably as a result of cell proliferation (see Methods), reaches a maximum size at about 9–11 DIV, and then declines by about 30% before reaching a stable size at about 17–22 DIV (ANOVA, $F = 4.53$ and $P < 0.01$). **e**, A transplanted interneuron precursor expressing cleaved caspase-3 (red) and β-actin:GFP (green; arrowhead) at 15 DAT. Scale bars, 50 μm (left) and 25 μm (right). **f**, Temporal profile of expression of cleaved caspase-3 in transplanted interneuron precursors. Cleaved caspase-3 is highest at 15 DAT, when the transplanted population reaches an intrinsic cellular age similar to that of endogenous interneurons during the peak of normal developmental cell death (Fig. 1b; ANOVA, $F = 17.79$ and $P < 0.0001$; $n = 5$ per time point). All error bars represent s.e.m.

during the peak of normal developmental cell death (Figs 1b and 2f). Taken together with the *in vitro* data (Fig. 2a–d), these findings suggest that interneuron cell death is timed by the intrinsic maturational state of the developing cells.

We also used heterochronic transplantation to introduce varying numbers of embryonic interneuron precursors into the neocortex. We expected that, if interneuron cell death were determined by intercellular competition for extrinsically derived signals, the amount of interneuron cell death would increase with larger transplant sizes. However, across initial transplant sizes that varied 200-fold (5×10^3, 5×10^4, 5×10^5 and 10^6 cells), similar fractions of the transplanted cells survived in the recipient neocortical hemisphere at 60 DAT ($20.8 \pm 2.4\%$, $22.2 \pm 1.4\%$, $17.8 \pm 0.6\%$ and $15.3 \pm 0.3\%$, respectively; Fig. 3a; ANOVA, $F = 0.34$ and $P = 0.12$). When 10^6 or 2×10^6 cells were transplanted, similar numbers of cells survived ($(1.65 \pm 0.18) \times 10^5$ cells versus $(1.53 \pm 0.01) \times 10^5$ cells, respectively; Student's *t*-test, $P = 0.58$), suggesting that the neocortical hemisphere has a limited capacity for about 1.6×10^5 additional interneurons. However, when the initial transplant size was far smaller than this theoretical limit, transplanted cell death still occurred, and it occurred at a constant rate. This finding indicates that interneuron cell death is not governed by competition for limited trophic signals derived from other cell types.

To further examine whether soluble neurotrophic signals regulate interneuron cell death, we studied the survival of mutant interneurons lacking the neurotrophin receptor, tropomyosin kinase receptor B (TrkB). We transplanted interneuron precursors from $TrkB^{-/-}$ donors[18] into P2 wild-type recipients and examined the survival of the cells at 60 DAT. The survival of transplanted $TrkB^{-/-}$ interneurons was similar to that of transplanted wild-type cells (Fig. 3b; $(2.32 \pm 0.32) \times 10^4$ wild-type cells versus $(2.20 \pm 0.20) \times 10^4$ $TrkB^{-/-}$ cells; Student's *t*-test, $P = 0.75$), indicating that the cell death of transplanted interneurons is not governed by neurotrophin signalling through TrkB. This finding is consistent with other reports suggesting that the death of neurons in the central nervous system during development is regulated by mechanisms other than neurotrophin signalling[6,19].

To confirm that transplanted interneuron cell death occurred through Bax-dependent apoptosis, we examined the survival of transplanted $Bax^{-/-}$ mutant cells[12], and compared their survival with that of transplanted wild-type and $Bax^{+/-}$ cells. We pooled counts of wild-type and $Bax^{+/-}$ interneurons, because endogenous interneuron cell death was not disrupted in P20 $Bax^{+/-}$ GAD67–GFP mutants ($(8.88 \pm 0.03) \times 10^5$ wild-type cells versus $(9.63 \pm 0.04) \times 10^5$ $Bax^{+/-}$ cells; Student's *t*-test, $P = 0.20$). At 60 DAT into P2 recipients, transplanted *Bax*-null interneurons survived in greater numbers than transplanted *Bax* heterozygous and wild-type interneurons (Fig. 3c; $(4.31 \pm 0.21) \times 10^4$ $Bax^{+/-}$ and wild-type cells versus $(9.11 \pm 1.63) \times 10^4$ wild-type cells; Student's *t*-test, $P = 0.03$), indicating that the death of transplanted interneurons, like that of endogenous interneurons, occurs at least partly through a Bax-dependent mechanism.

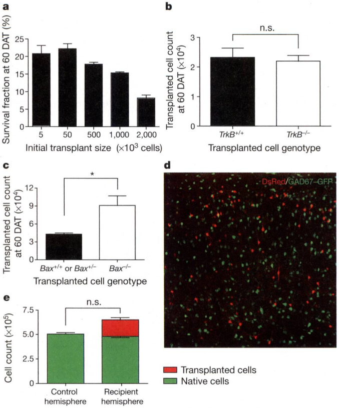


Figure 3 | Transplanted interneuron cell death is not governed by competition for survival signals derived from other cell types in the recipient neocortex. **a**, Over a broad range of transplant sizes (from 5×10^3 to 10^6 cells), nearly constant fractions of the transplanted populations survive at 60 DAT (about 15–22%; ANOVA, $F = 0.34$ and $P = 0.12$; $n = 6, 7, 3, 3$ per transplant size, respectively). When the initial transplant size is increased to 2×10^6 cells, a smaller fraction of transplanted cells survives in the recipient neocortex (about 8%; $n = 3$). **b**, Equal numbers of transplanted $TrkB^{-/-}$;β-actin:GFP interneurons and $TrkB^{+/+}$;β-actin:GFP interneurons survive in the recipient neocortex at 60 DAT (Student's *t*-test, $P = 0.75$; $n = 5$ per genotype). **c**, Transplanted cortical interneuron cell death occurs through a Bax-dependent mechanism. Greater numbers of $Bax^{-/-}$;β-actin:GFP cortical interneurons survive in the recipient cortex at 60 DAT, compared with transplanted $Bax^{+/+}$;β-actin:GFP and $Bax^{+/-}$;β-actin:GFP cortical interneurons (Student's *t*-test; $^*P < 0.05$; $n = 5$ for wild-type and $Bax^{+/-}$; $n = 6$ for $Bax^{-/-}$). **d**, Transplanted β-actin:DsRed interneurons (red) and endogenous GAD67–GFP neurons (green) at 60 DAT (initial transplant size 10^6 cells; scale bar, 150 μm). **e**, At 60 DAT, transplanted DsRed-labelled interneurons increase the cortical interneuron population size by 34% (red) without affecting the endogenous GAD67–GFP population (green; Student's *t*-test, $P = 0.28$; $n = 3$). n.s., not significant. All error bars represent s.e.m.

While our transplantation experiments strongly suggested that interneuron cell death is not determined through competition for extrinsic survival signals, it was possible that the transplanted cells competed with endogenous cells, and the survival of the transplanted interneurons occurred at the expense of endogenous interneuron survival. To examine this possibility, we transplanted 10^6 β-actin:DsRed MGE cells[20] to one neocortical hemisphere of P2 to P3 GAD67–GFP recipients, and then compared the number of endogenous interneurons between the recipient and contralateral control hemispheres (Fig. 3d). As expected (Fig. 3a), we observed an average of roughly 1.7×10^5 transplanted interneurons in the recipient cortical hemisphere at 60 DAT (Fig. 3e; mean $(1.69 \pm 0.41) \times 10^5$ cells). In the recipient and control hemispheres we observed equal numbers of endogenous interneurons (Fig. 3e; mean endogenous cell count, recipient hemisphere $= (4.81 \pm 0.12) \times 10^5$; mean endogenous cell count, control hemisphere $= (5.04 \pm 0.15) \times 10^5$; Student's *t*-test, $P = 0.28$), consistent with the findings presented in Supplementary Fig. 8, which indicated that transplantation did not affect the expression of cleaved caspase-3 in endogenous cells. The neocortex is thus able to support roughly 35% additional interneurons, with no effect on the endogenous interneuron population size. This suggests that developmental cell death does not tune the number of developing interneurons towards a cellular limit, as would occur if interneuron number were determined by the availability of limited, extrinsically derived survival signals.

Given that transplantation increases the number of interneurons in the neocortex, it offers a strategy for studying the relationship between interneuron number and cortical inhibition. To explore this relationship we transplanted varying numbers of interneuron precursors into P2 to P3 recipients, and then performed *in vitro* patch-clamp recordings on endogenous neocortical pyramidal neurons at 30–40 DAT. We recorded the amplitudes and frequencies of spontaneous inhibitory postsynaptic currents (sIPSCs; Fig. 4a) and then performed post-hoc quantification of transplanted interneuron cell densities.

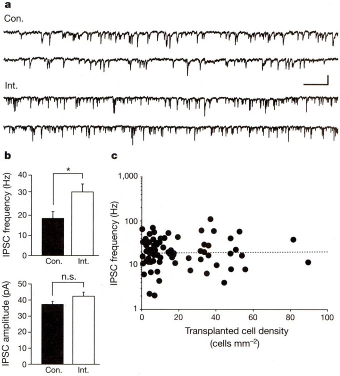


Figure 4 | Interneuron population size is not a primary determinant of the level of functional cortical inhibition. a, Representative traces of sIPSCs recorded from endogenous neocortical pyramidal neurons *in vitro* (top, control (medium vehicle (Con.); bottom, interneuron transplant recipient (Int)). Vertical scale bar, 40 pA; horizontal scale bar, 200 ms. **b,** Transplanted interneurons increase the frequency (top) but not the amplitude (bottom) of sIPSCs recorded at 30–40 DAT (Wilcoxon rank-sum test; $^*P < 0.05$ and $P = 0.22$, respectively; $n = 23$ recorded cells from control animals, $n = 37$ recorded cells from interneuron transplant recipients). The mean transplanted cell density for transplant recipient group was 23.3 ± 3.8 cells mm^{-2}. Error bars represent s.e.m. **c,** The frequency of sIPSCs onto host pyramidal neurons does not increase with the density of transplanted interneurons (linear regression analysis, slope = 0.0003, $r^2 = 0.0003$).

Consistent with previous findings[15,17], transplanted interneurons increased the frequency of sIPSCs onto endogenous pyramidal neurons (Fig. 4b; controls, 18.4 ± 3.4 Hz; transplant recipients, 31.7 ± 3.9 Hz; Wilcoxon rank-sum test, $P = 0.02$). However, the amplitudes of inhibitory events were not significantly increased by transplantation (Fig. 4b; controls, 37.3 ± 1.9 pA; transplant recipients, 42.4 ± 2.5 pA; Wilcoxon rank-sum test, $P = 0.22$). Inhibitory event frequencies did not increase with transplanted interneuron density (linear regression analysis, slope = 0.0003 and $r^2 = 0.0003$; Fig. 4c). Thus, the extent of cortical inhibition is more likely to be determined by mechanisms that adjust synaptic strength and number, rather than mechanisms that govern interneuron population size. These findings indicate that transplantation can add a limited amount of new inhibition to the neocortex, and this limit is reached with transplanted cell numbers much smaller than that which the neocortex can support.

Our findings suggest that interneuron cell death is regulated by intrinsically defined mechanisms. When interneuron precursors were cultured *in vitro* or heterochronically transplanted, they died when they reached a cellular age equivalent to that of endogenous interneurons during the peak of endogenous interneuron cell death (Figs 1 and 2). This suggests that interneuron cell death is timed by the expression of a maturational program intrinsic to interneurons, rather than the developmental state of the cortex itself. Similarly, the extent of interneuron cell death seems to be intrinsically defined: across a range of transplant sizes, a constant fraction of the transplanted interneurons died in the recipient cortex, even when the transplant size was significantly below the number of interneurons that the cortex could support (Fig. 3). Interneuron cell death is therefore unlikely to follow from intercellular competition for limiting survival signals derived from other cell types.

We propose two mechanisms that may govern the developmental cell death of cortical interneurons (Supplementary Fig. 1). In the first, which we refer to as 'cell-autonomous', interneuron cell death is intrinsically determined within each embryonic interneuron precursor. In this model, interneuron precursors would be individually destined to die in a manner independent from their interactions with other cell types. For example, the production of interneurons could occur with a certain rate of error[21] such that a fraction of defective interneuron precursors cannot survive past a certain cellular age. Similarly, a fixed fraction of interneuron precursors may be cell-autonomously programmed to die during a specific stage of their development. Alternatively, in a 'population-autonomous' mechanism, developing interneurons may require and compete for limiting survival signals produced by other isochronic interneurons. These neurotrophic signals, which may be obtained through cell–cell contact, synaptic transmission or neurotrophin signalling independent of TrkB, would be present in a quantity that scales to the number of isochronic developing interneurons. Either a cell-autonomous or population-autonomous mechanism could explain why cell death occurred at a constant rate across broad range of interneuron transplant sizes, and also why the survival of endogenous interneurons was not affected by the transplantation of additional interneurons.

Interneurons have a critical role in cortical physiology, and their dysfunction has been implicated in neurological disorders such as epilepsy, schizophrenia and Alzheimer's disease[22–24]. The detailed examination of interneuron cell death is thus expected to yield new insights into cortical development, the pathophysiology of brain disorders and the therapeutic application of neuronal transplantation.

METHODS SUMMARY

Cell transplantation. Interneuron precursor transplantation was performed as described previously[14,15,17]. For the transplantation of *TrkB* and *Bax* mutant cells (Fig. 3b, c), whole MGE explants were transplanted.

***In vitro* cell culture, immunostaining and cell counts.** As described previously[13], E13.5 GAD67–GFP MGE cells were added to feeder layers prepared from P0 to P2 CD1 mice. Immunostaining was then performed against the various markers at the specified time points. Cleaved caspase-3-positive cells and GAD67–GFP cells were counted under a standard fluorescence microscope.

***In vivo* immunostaining and cell counts.** Floating sections were immunostained against the specified marker(s). Cell counts were made throughout the entire depth, rostrocaudal extent and mediolateral extent of the neocortex; cell counts were not made in other cortical areas such as the olfactory bulb, piriform cortex or hippocampus. Cleaved caspase-3-positive cells and transplanted GFP interneuron precursors (initial transplant sizes of less than 10^5 cells) were directly counted under a standard fluorescence microscope. To quantify larger populations (endogenous GAD67–GFP cells and initial transplant sizes of 5×10^5 cells or more), design-based stereology was performed with StereoInvestigator (MicroBrightField).

Electrophysiology and cell counts. sIPSCs were recorded from layer 2/3 pyramidal cells in coronal slices prepared from recipient somatosensory cortex. All electrophysiology was performed with the experimenter blinded to the number of transplanted interneurons.

Statistical analysis. An analysis of variance was used to test for differences among three or more groups. A Student's t-test was used to compare cell counts between two groups. A Wilcoxon rank-sum test and a linear regression analysis were used to analyse the sIPSC data. Statistical analyses were performed with Prism 4.0 (Graphpad) and Sigma Plot 12 (Systat Software).

Full Methods and any associated references are available in the online version of the paper.

Received 25 June 2009; accepted 17 August 2012.

Published online 7 October 2012.

1. Buss, R. R., Sun, W. & Oppenheim, R. W. Adaptive roles of programmed cell death during nervous system development. *Annu. Rev. Neurosci.* **29,** 1–36 (2006).
2. Oppenheim, R. W. Cell death during development of the nervous system. *Annu. Rev. Neurosci.* **14,** 453–501 (1991).

3. Hamburger, V. & Levi-Montalcini, R. Proliferation, differentiation and degeneration in the spinal ganglia of the chick embryo under normal and experimental conditions. *J. Exp. Zool.* **111,** 457–501 (1949).
4. Levi-Montalcini, R. The development of the acoustico-vestibular centers in the chick embryo in the absence of the afferent root fibers and of descending fiber tracts. *J. Comp. Neurol.* **91,** 209–241 (1949).
5. Boya, P. & De la Rosa, E. J. Cell death in early neural life. *Birth Defects Res.* **75,** 281–293 (2005).
6. Nikoletopoulou, V. *et al.* Neurotrophin receptors TrkA and TrkC cause neuronal death whereas TrkB does not. *Nature* **467,** 59–64 (2010).
7. Polleux, F., Whitford, K. L., Dijkhuizen, P. A., Vitalis, T. & Ghosh, A. Control of cortical interneuron migration by neurotrophins and PI3-kinase signaling. *Development* **129,** 3147–3160 (2002).
8. Gorba, T. & Wahle, P. Expression of TrkB and TrkC but not BDNF mRNA in neurochemically identified interneurons in rat visual cortex *in vivo* and in organotypic cultures. *Eur. J. Neurosci.* **11,** 1179–1190 (1999).
9. Tamamaki, N. *et al.* Green fluorescent protein expression and colocalization with calretinin, parvalbumin, and somatostatin in the GAD67–GFP knock-in mouse. *J. Comp. Neurol.* **467,** 60–79 (2003).
10. Wonders, C. P. & Anderson, S. A. The origin and specification of cortical interneurons. *Nature Rev. Neurosci.* **7,** 687–696 (2006).
11. Sahara, S., Yanagawa, Y., O'Leary, D. D. M. & Stevens, C. F. The fraction of cortical GABAergic neurons is constant from near the start of cortical neurogensis to adulthood. *J. Neurosci.* **32,** 4755–4761 (2012).
12. Knudson, C. M., Tung, K. S., Tourtellote, W. G., Brown, G. A. & Korsmeyer, S. J. Bax-deficient mice with lymphoid hyperplasia and male germ cell death. *Science* **270,** 96–99 (1995).
13. Xu, Q., Cobos, I., De la Cruz, E., Rubenstein, J. L. & Anderson, S. A. Origins of cortical interneuron subtypes. *J. Neurosci.* **24,** 2612–2622 (2004).
14. Wichterle, H., Garcia-Verdugo, J. M., Herrera, D. G. & Alvarez-Buylla, A. Young neurons from medial ganglionic eminence disperse in adult and embryonic brain. *Nature Neurosci.* **2,** 461–466 (1999).
15. Alvarez-Dolado, M. *et al.* Cortical inhibition modified by embryonic neural precursors grafted into the postnatal brain. *J. Neurosci.* **26,** 7380–7389 (2005).
16. Hadjantonakis, A., Gertenstein, M., Ikawa, M., Okabe, M. & Nagy, A. Generating green fluorescent mice by germline transmission of green fluorescent ES cells. *Mech. Dev.* **76,** 79–90 (1998).
17. Southwell, D. G., Froemke, R. C., Alvarez-Buylla, A., Stryker, M. P. & Gandhi, S. P. Cortical plasticity induced by inhibitory neuron transplantation. *Science* **327,** 1145–1148 (2010).
18. Galvao, R. P., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Brain-derived neurotrophic factor signaling does not stimulate subventricular zone neurogenesis in adult mice and rats. *J. Neurosci.* **28,** 13368–13383 (2008).
19. Silos-Santiago, I. *et al.* Severe sensory deficits but normal CNS development of newborn mice lacking TrkB and TrkC tyrosine protein kinase receptors. *Eur. J. Neurosci.* **9,** 2045–2056 (1997).
20. Vintersten, K. *et al.* Mouse in red: red fluorescent protein expression in mouse ES cells, embryos, and adult animals. *Genesis* **40,** 241–246 (2004).
21. Saunders, J. Death in embryonic systems. *Science* **154,** 604–612 (1966).
22. Cossart, R., Bernard, C. & Ben-Ari, Y. Multiple facets of GABAergic neurons and synapses: multiple fates of GABA signaling in epilepsies. *Trends Neurosci.* **28,** 108–115 (2005).
23. Marin, O. Interneuron dysfunction in psychiatric disorders. *Nature Rev. Neurosci.* **13,** 107–120 (2012).
24. Verret, L. *et al.* Inhibitory interneuron deficit links altered network activity and cognitive dysfunction in Alzheimer model. *Cell* **149,** 708–721 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank Y. Yanagawa for GAD67–GFP mice, L. Reichardt for *TrkB* mutant mice and R. Romero for technical contributions. D.G.S. and M.F.P. were supported by training grants from the California Institute for Regenerative Medicine. J.Y.S. was supported by a fellowship from the National Institute of Neurologic Disorders and Stroke (F32NS061497). This work was supported by the California Institute for Regenerative Medicine (A.A.-B., TR2-01749), the John G. Bowes Research Fund (A.A.-B.), the Spanish Ministry of Science and Innovation (J.M.G-V., SAF-2008-01274), and the National Institutes of Health (J.L.R., S.C.B. and A.A.-B., R01 NS071785; A.A.-B., R01 NS048528).

Author Contributions D.G.S. and A.A.-B. devised experiments. D.G.S. performed all experiments, with the exception of the following: M.F.P. performed *in vitro* culture experiments and measurements of Ki67 labelling, R.P.G. collected cleaved caspase-3 counts in GAD67–GFP mice, C.A.-C. and J.M.G-V. produced electron micrographs, Y.T. performed cell transplantations for the examination of Ki67 labelling, and D.L.J., J.Y.S. and R.F. performed electrophysiological recordings. D.G.S. and A.A.-B. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.G.S. (dereksouthwell@gmail.com) or A.A.-B. (abuylla@stemcell.ucsf.edu).

doi:10.1038/nature11537

In vivo genome editing using a high-efficiency TALEN system

Victoria M. Bedell[1]*, Ying Wang[2]*, Jarryd M. Campbell[1]*, Tanya L. Poshusta[1], Colby G. Starker[3], Randall G. Krug II[1], Wenfang Tan[3], Sumedha G. Penheiter[1], Alvin C. Ma[1,4], Anskar Y. H. Leung[4], Scott C. Fahrenkrug[3,5], Daniel F. Carlson[3,5], Daniel F. Voytas[3], Karl J. Clark[1], Jeffrey J. Essner[2] & Stephen C. Ekker[1]

The zebrafish (*Danio rerio*) is increasingly being used to study basic vertebrate biology and human disease with a rich array of *in vivo* genetic and molecular tools. However, the inability to readily modify the genome in a targeted fashion has been a bottleneck in the field. Here we show that improvements in artificial transcription activator-like effector nucleases (TALENs) provide a powerful new approach for targeted zebrafish genome editing and functional genomic applications[1–5]. Using the GoldyTALEN modified scaffold and zebrafish delivery system, we show that this enhanced TALEN toolkit has a high efficiency in inducing locus-specific DNA breaks in somatic and germline tissues. At some loci, this efficacy approaches 100%, including biallelic conversion in somatic tissues that mimics phenotypes seen using morpholino-based targeted gene knockdowns[6]. With this updated TALEN system, we successfully used single-stranded DNA oligonucleotides to precisely modify sequences at predefined locations in the zebrafish genome through homology-directed repair, including the introduction of a custom-designed *EcoRV* site and a modified *loxP* (*mloxP*) sequence into somatic tissue *in vivo*. We further show successful germline transmission of both *EcoRV* and *mloxP* engineered chromosomes. This combined approach offers the potential to model genetic variation as well as to generate targeted conditional alleles.

Custom zinc finger nucleases (ZFNs)[7–9] and TALENs[1–5] have been used to introduce locus-specific double-stranded breaks in the zebrafish genome, generating dozens of mutant alleles[10]. Recent work has been facilitated by the relatively straightforward DNA base recognition cipher underlying TALEN technology[11,12]. However, the efficacy of previously described custom sequence-specific nucleases was limiting in some applications[1–5,7–9]. For example, standard TALENs using the pTAL scaffold[13] (Supplementary Fig. 1) targeting exon 2 of the zebrafish *ponzr1* locus[14] resulted in a measurable level of locus modification in somatic tissue (median value of 5%; Fig. 1b, c). This pTAL-*ponzr1* pair yielded 4 germline-transmitting founder animals carrying a mutation in *ponzr1* out of the 24 tested (Supplementary Fig. 3d). TALENs against a second locus (*crhr1*) using the pTAL scaffold yielded a modest rate of locus modification (<1%; Fig. 1b, c). These results are characteristic of the standard TALEN efficacy range, demonstrating room for improvement.

Multiple TALEN scaffold designs have been described[13,15,16], including those with different amino- and carboxy-terminal truncations, diverse FokI nuclease linkers, and various nuclear localization sequences. To improve *in vivo* efficacy, we tested the GoldyTALEN scaffold (Supplementary Fig. 1 and Supplementary Fig. 2) in a messenger RNA expression vector backbone (pT3TS[17]) using DNA analysis that measures the loss of a restriction enzyme recognition sequence at the TALEN cut site (Fig. 1a). Using the same recognition domains in the GoldyTALEN scaffold, there is a sixfold increase in somatic gene modification at the *ponzr1* locus (Fig. 1b, c and Supplementary Fig. 3b) over the pTAL scaffold. The germline modification rate was similarly increased when switching scaffolds, from 17% (4/24; pTAL-*ponzr1*; Supplementary Fig. 3d) to 71% (10/14; GoldyTALEN-*ponzr1*; Supplementary Fig. 3e). We also detected improved efficacy using a cell-free assay system with *in vitro*-translated TALEN protein and purified *ponzr1* PCR DNA (Fig. 1d). The GoldyTALENs against *crhr1* showed an increase in the genome modification rate, improving from <1% to 7% median cutting efficacy (Fig. 1b, c and Supplementary Fig. 3c). Sequence comparisons of pTAL and GoldyTALEN scaffolds in both loci demonstrate similar insertions or deletions (indels) at the cut site, which is diagnostic of non-homologous end joining (NHEJ) repair (Supplementary Fig. 3).

To further test the efficacy of the GoldyTALEN scaffold, we generated TALENs against three additional loci (*moesina*, also known as *msna*, *ppp1cab* and *cdh5*; Supplementary Fig. 4a). We observed efficient gene modification at each locus (5 out of 5 loci in total; Fig. 1 and Fig. 2a). In three instances, the mutagenesis efficiency ranged from 70 to 100% as demonstrated by loss of the restriction enzyme recognition sequence at the TALEN cut sites (Fig. 2a) and DNA sequence analyses (Supplementary Fig. 4b–d) of amplicons from pooled injected embryos. To determine the time course of the GoldyTALEN-induced changes, we examined restriction enzyme nuclease activity at 256-cell, 28 h post-fertilization (hpf) and 50 hpf stages. A majority of the DNA was modified by the 256-cell stage (Supplementary Fig. 5). Together, these results indicate early, efficient gene targeting in somatic tissues, including biallelic conversion in some animals. Somatic targeting efficacy using the GoldyTALEN scaffold compares favourably with previous TALEN scaffolds in zebrafish, with three out of five GoldyTALENs demonstrating as high or higher mutation frequency as any of the previously reported loci using the first generation TALEN systems[1–5].

In response to the increased efficacy of the GoldyTALENs, we investigated whether injection of TALENs could recapitulate a known morpholino[6] loss of function phenotype. We conducted a dose–response curve of the *moesina*, *ppp1cab* and *cdh5* GoldyTALEN pairs, optimizing GoldyTALEN concentration to the number of embryos with biallelic changes, and per cent dead or malformed embryos (Supplementary Fig. 6). Embryos injected with either *cdh5* GoldyTALENs (Fig. 2d) or morpholinos[18] (Fig. 2c) showed similar vascular phenotypes: pronounced cardiac oedema (Fig. 2b, top panels), loss of patent lumens in the Tg(*fli1-egfp*)^{y1} vasculature[19] (Fig. 2c, d, bottom panels), and loss of circulating Tg(*gata1:dsred*)^{sd2} red blood cells[20] (Fig. 2c, d, bottom panels, and Supplementary Movies 1–3). A similar pericardial oedema phenotype was observed in F1 offspring from F0 *cdh5* founder incrosses (data not shown), suggesting specificity of the phenotype described in F0 fish to *cdh5* loss of function. Furthermore, *cdh5* GoldyTALEN-injected embryos have little or no Cdh5 protein

---

[1] Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, Minnesota 55905, USA. [2] Department of Genetics, Development and Cell Biology, Iowa State University, Ames, Iowa 50011, USA. [3] Department of Genetics, Cell Biology and Development, the Center for Genome Engineering, and the Department of Animal Science, University of Minnesota, Minneapolis, Minnesota 55455, USA. [4] Department of Medicine, LKS Faculty of Medicine, the University of Hong Kong, Pok Fu Lam Road, Hong Kong. [5] Recombinetics, Inc., 2575 University Ave. West, Suite 100, Saint Paul, Minnesota 55114, USA.
*These authors contributed equally to this work.

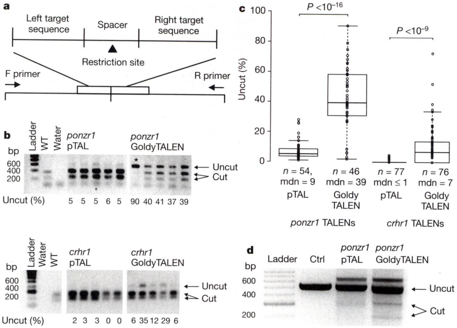


Figure 1 | Second-generation GoldyTALEN scaffold improves genome-editing efficacy. a, Schematic showing the layout of TALEN target sites. TALENs were targeted to flanking sequences surrounding a restriction enzyme site for easy screening through introduction of a restriction fragment length polymorphism. **b**, Relative activity of the GoldyTALEN and pTAL scaffolds at two loci, *ponzr1* and *crhr1*. Under each lane is the percent uncut DNA of a single larva, illustrating the increased activity of GoldyTALEN. WT, wild type. **c**, Whisker plots of the percent uncut DNA demonstrates TALEN cutting efficiency at two loci. *ponzr1* TALENs demonstrate a significant ($P < 10^{-16}$), sixfold increase in activity using GoldyTALEN. *crhr1* TALENs also demonstrate a significant ($P < 10^{-9}$), 15-fold increase in activity. *n*, number of embryos screened; mdn, the median percent cut. **d**, The *ponzr1* GoldyTALENs were more active in a cell-free restriction enzyme digestion assay. *ponzr1* DNA is labelled in both uncut and cut forms. Ctrl, negative control.

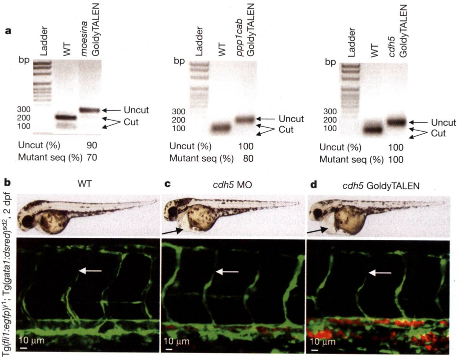

a
Ladder
WT
moesina GoldyTALEN
bp
300
200
100
Uncut
Cut
Uncut (%) 90
Mutant seq (%) 70
Ladder
WT
ppp1cab GoldyTALEN
bp
300
200
100
Uncut
Cut
Uncut (%) 100
Mutant seq (%) 80
Ladder
WT
cdh5 GoldyTALEN
bp
300
200
100
Uncut
Cut
Uncut (%) 100
Mutant seq (%) 100
b
WT
c
cdh5 MO
d
cdh5 GoldyTALEN
Tg(fli1:egfp)^y1; Tg(gata1:dsred)^sd2, 2 dpf
10 μm
10 μm
10 μm

Figure 2 | Increased TALEN efficiency results in biallelic gene targeting. a, GoldyTALENs were designed against the *moesina*, *ppp1cab* and *cdh5* genes. All three gene targets contained a restriction enzyme site within the spacer region between the TALEN binding sites. Injection of GoldyTALEN mRNAs demonstrated a nearly complete loss of the restriction enzyme site in the amplicons of somatic tissue. Each lane is the amplification product from a group of 10 embryos. Mutant seq (%), percentage of amplicons that carry mutant sequences as determined by sequencing 10 clones (Supplementary Fig. 4). **b–d**, Injection of *cdh5* GoldyTALENs (**d**) phenocopies the morpholino-based loss-of-function phenotype (**c**). Bright field images (top panels) show pronounced cardiac oedema (arrows) in both GoldyTALEN-injected (**d**) and morpholino-injected (**c**) larvae at 2 days post fertilization (dpf). Using the Tg(*fli1-egfp*)^{y1} line, the intersomitic vessels were visualized (bottom panels) and show a loss of lumen formation (white arrow) in both the morpholino-injected (**c**) and GoldyTALEN-injected larvae (**d**). The Tg(*gata1:dsred*)^{sd2} line revealed reduced circulation in GoldyTALEN- and morpholino-injected larvae, demonstrated by the increase in red fluorescence in the confocal images (see Supplementary Movies 1–3).

(Supplementary Fig. 7). Together, these results indicate that the GoldyTALEN platform can achieve efficient biallelic targeting recapitulating known loss-of-function phenotypes. Furthermore, these data demonstrate that GoldyTALENs have the potential to be a complementary, but distinct, approach to morpholino-based somatic phenotype assessment.

The biallelic GoldyTALEN-injected fish were raised to assess germline mutation transmission. The *moesina*, *ppp1cab* or *cdh5* F0 founders were outcrossed. Ten pooled F1 embryos were screened and showed a 9 to 55% locus mutation frequency (Supplementary Fig. 8a–c). From two founder F0 outcrosses per locus, 10 individual F1 embryos were sequenced with mutant alleles identified in 20% to 100% of the F1 offspring (Supplementary Fig. 8). Furthermore, in two out of three of these loci we detected germline mosaicism, indicating several independent repair events. These data indicate that the efficient somatic TALEN targeting is effectively passed through the germline.

Recent *in vitro* work demonstrates that single-stranded DNA (ssDNA) can be an effective donor for homology-directed repair (HDR)-based genome editing at a ZFN-induced double-stranded break[21,22]. With the highly efficient genome modification success of GoldyTALENs, we hypothesized that synthetic oligonucleotides designed to span the predicted TALEN cut site could serve as a template for HDR *in vivo* (Fig. 3). Using *ponzr1* as a test locus, we introduced an EcoRV restriction site by co-injection of *ponzr1* GoldyTALENs and a ssDNA oligonucleotide (Fig. 3a). In these experiments, 42 out of 74 injected embryos had a detectable level of chromosomes containing the introduced *EcoRV* sequence with an estimated 9% ratio of converted chromosomes in these animals (Supplementary Fig. 9a). Sequence analysis indicated two precisely modified chromosome events from different larvae (Supplementary Fig. 9b) demonstrating successful somatic HDR at the *ponzr1* locus. Other events show precise addition at the 3′ end while small indels were noted at the 5′ side of the modification site (Supplementary Fig. 9b). Several homology arm lengths were tested for the highest HDR signal. In this experimental approach, an increase in homology arm length that spanned the TALEN binding site decreased the frequency of HDR events (Supplementary Table 1).

To test whether the HDR sequence modification was stably maintained in zebrafish somatic tissue, fin biopsies from 2-month-old fish were assayed for addition of the *EcoRV* sequence at the *ponzr1* locus. Out of 186 fish, 8 showed a visible incorporation of *EcoRV* (Supplementary Fig. 9c). To determine whether a lack of somatic *EcoRV* incorporation also indicated a lack of germline incorporation, 13 randomly selected fish with *EcoRV*-negative fin biopsies were outcrossed. The offspring from all 13 adults were negative for *EcoRV* incorporation at the *ponzr1* locus (clutch sizes ranged from 16 to 96 embryos). Therefore, fin-biopsy-positive fish were prioritized for determining germline transmission. Outcross embryos from three out of four fin-tissue-positive fish yielded clutches with introduction of the *EcoRV* site at the *ponzr1* locus (Fig. 3b). Two out of three of these germline fish demonstrated precise *EcoRV* addition (Fig. 3c).

We next asked whether TALEN/oligonucleotide co-injection could introduce larger sequences such as a *loxP* site, an essential step in making Cre-dependent conditional genetic alleles. We used TALENs against an intron in the *crhr2* gene and a ssDNA oligonucleotide were used to add a modified *loxP*$^{\mathrm{JTZ17}}$ (*mloxP*)[23] site at this location (Fig. 4a). PCR analysis demonstrates somatic introduction of the *mloxP* sequence at the *crhr2* TALEN cut site (Supplementary Fig. 10a). Sequence characterization confirmed integration of the *mloxP* site in three assayed somatic chromosomes (Supplementary Fig. 10b). A similar method was used to introduce an *mloxP* sequence at the *ponzr1* locus (Supplementary Fig. 11a). Sequencing confirmed precise somatic addition at this locus (Supplementary Fig. 11b, c).

Maintenance of somatic *mloxP*-modified *crhr2* chromosomes by fin biopsy was used to identify germline transmission of the *mloxP* sequence. Positive chromosomes were detected by quantitative PCR in 20 out of 53 animals (Supplementary Fig. 10a). Embryos were obtained from 16 of the somatic-positive fish as well as 42 fish that had not been pre-screened by PCR. Both groups transmitted HDR events through the germline (Fig. 4b). However, no significant enrichment for probable germline transmitting animals was noted, perhaps owing to the less stringent PCR assay than that used for *ponzr1*. In total, 6 out of 58 injected animals transmitted *mloxP*-modified chromosomes through the germline at the *crhr2* locus (Fig. 4b). Sequence confirmation of three of these fish demonstrated a precise HDR event as well as other, non-precise events (Fig. 4c).

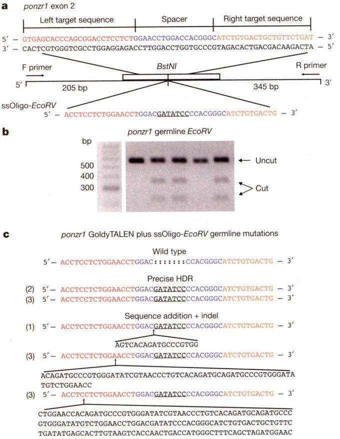

a
ponzr1 exon 2
Left target sequence
Spacer
Right target sequence
5′ – GTGAGCACCCAGCGGACCTCCTCTGGAACCTGGACCACGGGCATCTGTGACTGCTGTTCTGAT – 3′
3′ – CACTCGTGGGTCGCCTGGAGGAGACCTTGGACCTGGTGCCCGTAGACACTGACGACAAGACTA – 5′
F primer
BstNI
R primer
5′
205 bp
345 bp
3′
ssOligo-EcoRV
5′ – ACCTCCTCTGGAACCTGGACGATATCCCCACGGGCATCTGTGACTG – 3′
b
ponzr1 germline EcoRV
bp
500
400
300
Uncut
Cut
c
ponzr1 GoldyTALEN plus ssOligo-EcoRV germline mutations
Wild type
5′– ACCTCCTCTGGAACCTGGAC:::::::CCACGGGCATCTGTGACTG – 3′
Precise HDR
(2) 5′– ACCTCCTCTGGAACCTGGACGATATCCCCACGGGCATCTGTGACTG – 3′
(3) 5′– ACCTCCTCTGGAACCTGGACGATATCCCCACGGGCATCTGTGACTG – 3′
Sequence addition + indel
(1) 5′– ACCTCCTCTGGAACCTGGACGATATCCCCACGGGCATCTGTGACTG – 3′
AGTCACAGATGCCCGTGG
(3) 5′– ACCTCCTCTGGAACCTGGACGATATCCCCACGGGCATCTGTGACTG – 3′
ACAGATGCCCGTGGGATATCGTAACCCTGTCACAGATGCAGATGCCCGTGGGATA
TGTCTGGAACC
(3) 5′– ACCTCCTCTGGAACCTGGACGATATCCCCACGGGCATCTGTGACTG – 3′
CTGGAACCACAGATGCCCGTGGGATATCGTAACCCTGTCACAGATGCAGATGCCC
GTGGGATATGTCTGGAACCTGGACGATATCCCACGGGCATCTGTGACTGCTGTTC
TGATATGAGCACTTGTAAGTCACCAACTGACCATGGGCTTTCAGCTAGATGGAAC

Figure 3 | Targeted genome editing using GoldyTALENs. **a**, A schematic of the *ponzr1* locus with the ssDNA sequence used to introduce a targeted exogenous *EcoRV* sequence (underlined, note the extra C added to make the sequence mutagenic) into the genome *in vivo*. The left and right TALEN binding sites are shown in red and orange, respectively, and the spacer region is in blue. **b**, A representative gel from founder fish no. 2 demonstrating germline transmission of the HDR-based *EcoRV* sequence. Three out of four fin-tissue-positive fish demonstrated germline transmission of the *EcoRV* sequence. **c**, Sequence analysis of the three germline-transmitting lines. The first fish transmitting HDR-based genome changes through the germline (1) yielded 7 out of 96 embryos with an incorporated *EcoRV* site. The genomes of all 7 embryos showed the same modified sequence. The second founder fish (2) yielded 7 out of 46 embryos with *EcoRV* incorporation. All 7 embryos showed precise HDR-based addition of the *EcoRV* sequence. The third fish with germline transmission (3) yielded 5 out of 18 embryos with an incorporated *EcoRV* site, and showed a mosaic germline as demonstrated by offspring with three different modified sequences. One embryo included precise HDR-based *EcoRV* addition. The other 4 embryos contained sequence insertions on the 5′ end with two embryos each harbouring the specific sequences changes.

Here, we focused on local genome-editing changes induced by TALENs, especially those induced by HDR. However, more complete analyses will be required to assess any off-target effects of TALENs or ssDNA-based HDR. Whole-genome sequencing on germline-transmitting fish from different parental lines would be particularly instructive. Should this analysis demonstrate off-target mutations, TALENs using obligate heterodimer-based nuclease fusions have

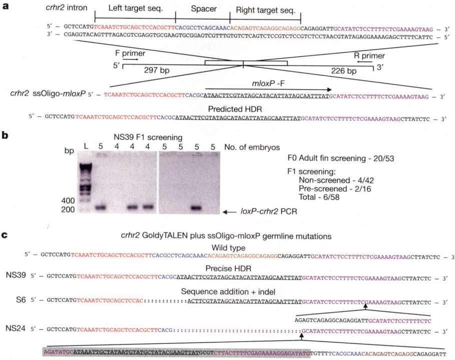

a
crhr2 intron
Left target seq.
Spacer
Right target seq.
5′ – GCTCCATGTCAAATCTGCAGCTCCACGCTTCACGCCTCAGCAAACACAGAGTCAGAGGCAGAGGCAGAGGATTGCATATCTCCTTTTCTCGAAAAGTAAG – 3′
3′ – CGAGGTACAGTTTAGACGTCGAGGTGCGAAGTGCGGAGTCGTTTGTGTCTCAGTCTCCGTCTCCGTCTCCTAACGTATAGAGGAAAAGAGCTTTTCATTC – 5′
F primer
5′
297 bp
R primer
3′
226 bp
mloxP -F
crhr2 ssOligo-mloxP 5′ – TCAAATCTGCAGCTCCACGCTTCACGCATAACTTCGTATAGCATACATTATAGCAATTTATGCATATCTCCTTTTCTCGAAAAGTAAG – 3′
Predicted HDR
5′ – GCTCCATGTCAAATCTGCAGCTCCACGCTTCACGCATAACTTCGTATAGCATACATTATAGCAATTTATGCATATCTCCTTTTCTCGAAAAGTAAGCTTATCTC – 3′
b
NS39 F1 screening
bp
L 5 4 4 4 5 5 5 5
No. of embryos
400
200
F0 Adult fin screening - 20/53
F1 screening:
Non-screened - 4/42
Pre-screened - 2/16
Total - 6/58
loxP-crhr2 PCR
c
crhr2 GoldyTALEN plus ssOligo-mloxP germline mutations
Wild type
5′ – GCTCCATGTCAAATCTGCAGCTCCACGCTTCACGCCTCAGCAAACACAGAGTCAGAGGCAGAGGCAGAGGATTGCATATCTCCTTTTCTCGAAAAGTAAGCTTATCTC – 3′
Precise HDR
NS39 5′ – GCTCCATGTCAAATCTGCAGCTCCACGCTTCACGCATAACTTCGTATAGCATACATTATAGCAATTTATGCATATCTCCTTTTCTCGAAAAGTAAGCTTATCTC – 3′
Sequence addition + indel
S6 5′ – GCTCCATGTCAAATCTGCAGCTCCAC:::::::::::ACTTCGTATAGCATACATTATAGCAATTTATGCATATCTCCTTTTCTCGAAAAGTAAGCTTATCTC – 3′
AGAGTCAGAGGCAGAGGATTGCATATCTCCTTTTCTC
NS24 5′ – GCTCCATGTCAAATCTGCAGCTCCACGCTTCACG:::::::::::::::::::::::::::::::::::::::GCATATCTCCTTTTCTCGAAAAGTAAGCTTATCTC – 3′
AGATATGCATAAATTGCTATAATGTATGCTATACGAAGTTATGCGTCTTACTTTTCGAGAAAAGGAGATATGTGTTTTCACGCAAACACAGAGTCAGAGGCAGAGGATT

Figure 4 | Germline *mloxP* integration into the *crhr2* locus. **a**, A diagram of the TALEN target sites with the *mloxP* ssDNA oligonucleotide. The left and right TALEN target sequences are red and orange, respectively, the spacer region is blue, and the right homology arm of the oligonucleotide is in purple. The *mloxP* sequence is underlined. **b**, Germline screening of the *crhr2* locus. 53 adult fish were pre-screened via fin biopsy. Of those pre-screened, 20 demonstrated mloxP maintenance. A total of 16 F0 fish were outcrossed with 2 showing germline transmission. A total of 42 unscreened F0 fish were outcrossed and 4 demonstrated germline transmission. **c**, Sequence confirmation of three *mloxP* germline fish. One fish demonstrated precise germline HDR whereas two showed indels. In NS24, the reverse complement of the *mloxP* was noted (shaded in grey).

recently been reported as an alternative approach[3,5,24]. Using obligate heterodimers in the GoldyTALEN scaffold is one future method for potentially optimizing HDR-directed gene-editing specificity.

To our knowledge, these results represent the first description of successful HDR in zebrafish and the first demonstration of HDR using ssDNA as a donor template *in vivo*. This approach complements the error-prone NHEJ toolkit for model organisms (Fig. 5). The use of ssDNA facilitates an array of genome changes, including the introduction of single-nucleotide polymorphisms for vertebrate genetic applications. The asymmetry in precise editing suggests an additional mechanism for genome editing that incorporates both HDR and NHEJ (Fig. 5). For example, the donor ssDNA may serve as a primer for new strand synthesis at the TALEN break. Extension from the 3′ end of the oligonucleotide would create long regions of homology for recombination. However, the 5′ end of the oligonucleotide limits the extent of strand invasion and a limited opportunity for HDR. This leads to 5′ end resolution by either HDR or NHEJ. For applications where new sequences are introduced into non-coding genomic regions, such as the introduction of *loxP* sites into intronic sequences, either event will probably be of high utility.

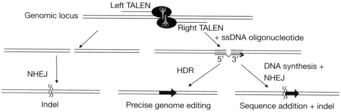

Left TALEN
Genomic locus
Right TALEN
+ ssDNA oligonucleotide
5′
3′
NHEJ
HDR
DNA synthesis + NHEJ
Indel
Precise genome editing
Sequence addition + indel

Figure 5 | *In vivo* TALEN-induced genome editing outcomes. TALENs efficiently create double-stranded breaks in chromosomal DNA and catalyse three major outcome classes. First, error-prone NHEJ produces an indel in and near the spacer region of the TALEN binding site. If a complementary ssDNA oligonucleotide is also added, two different outcomes are noted. First, HDR precisely uses the exogenous sequence information in the ssDNA to add sequence at the cut site. Alternatively, ssDNA acts as a primer for 3′ integration of the oligonucleotide but the 5′ end undergoes error-prone NHEJ[22].

Using the zebrafish, we report an updated TALEN system for use in genome modification and functional genomic applications. The high efficacy enables new approaches, including somatic gene targeting for reverse genetics applications. Furthermore, we show that synthetic ssDNA oligonucleotides can be used with this TALEN system for genome editing, including the precise introduction of exogenous DNA sequence at a specific locus. Although deployed here in zebrafish, this approach has the potential to be effective for *in vivo* applications in a wide array of model organisms.

METHODS SUMMARY

TALENs were assembled via the GoldenGate method[13]. For ease of analysis, TALEN recognition sequences flanked a unique restriction site within the targeted gene. TALEN repeat variable di-residues (RVDs) were cloned into a pT3TS[17]-driven TALEN scaffold, and mRNA was injected into single-cell zebrafish embryos. The injected larvae were either molecularly tested or raised for germline mutation analysis. Somatic and germline TALEN-induced mutations were evaluated via PCR and restriction fragment length polymorphisms. To induce HDR events, singe-stranded DNA oligonucleotides with either an *EcoRV* or *mloxP* site were designed with short homology arms around a TALEN target site and were injected into one-cell zebrafish embryos. PCR analysis of modified loci was used to detect the resulting somatic and germline HDR events.

Full Methods and any associated references are available in the online version of the paper.

Received 3 May; accepted 6 September 2012.

Published online 23 September 2012.

1. Huang, P. *et al.* Heritable gene targeting in zebrafish using customized TALENs. *Nature Biotechnol.* **29,** 699–700 (2011).
2. Sander, J. D. *et al.* Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nature Biotechnol.* **29,** 697–698 (2011).

3. Cade, L. *et al.* Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res.* **40,** 8001–8010 (2012).
4. Moore, F. E. *et al.* Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). *PLoS ONE* **7,** e37877 (2012).
5. Dahlem, T. J. *et al.* Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet.* **8,** e1002861 (2012).
6. Nasevicius, A. & Ekker, S. C. Effective targeted gene 'knockdown' in zebrafish. *Nature Genet.* **26,** 216–220 (2000).
7. Doyon, Y. *et al.* Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nature Biotechnol.* **26,** 702–708 (2008).
8. Meng, X., Noyes, M. B., Zhu, L. J., Lawson, N. D. & Wolfe, S. A. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nature Biotechnol.* **26,** 695–701 (2008).
9. Foley, J. E. *et al.* Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN). *PLoS ONE* **4,** e4348 (2009).
10. Lawson, N. D. & Wolfe, S. A. Forward and reverse genetic approaches for the analysis of vertebrate development in the zebrafish. *Dev. Cell* **21,** 48–64 (2011).
11. Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326,** 1501 (2009).
12. Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326,** 1509–1512 (2009).
13. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39,** e82 (2011).
14. Bedell, V. M. *et al.* The lineage-specific gene *ponzr1* is essential for zebrafish pronephric and pharyngeal arch development. *Development* **139,** 793–804 (2012).
15. Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nature Biotechnol.* **29,** 143–148 (2011).
16. Mussolino, C. *et al.* A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* **39,** 9283–9293 (2011).
17. Hyatt, T. M. & Ekker, S. C. Vectors and techniques for ectopic gene expression in zebrafish. *Methods Cell Biol.* **59,** 117–126 (1998).
18. Wang, Y. *et al.* Moesin1 and Ve-cadherin are required in endothelial cells during *in vivo* tubulogenesis. *Development* **137,** 3119–3128 (2010).
19. Lawson, N. D. & Weinstein, B. M. *In vivo* imaging of embryonic vascular development using transgenic zebrafish. *Dev. Biol.* **248,** 307–318 (2002).
20. Traver, D. *et al.* Transplantation and *in vivo* imaging of multilineage engraftment in zebrafish bloodless mutants. *Nature Immunol.* **4,** 1238–1246 (2003).
21. Chen, F. *et al.* High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nature Methods* **8,** 753–755 (2011).
22. Radecke, S., Radecke, F., Cathomen, T. & Schwarz, K. Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications. *Mol. Ther.* **18,** 743–753 (2010).
23. Thomson, J. G., Rucker, E. B. III & Piedrahita, J. A. Mutational analysis of *loxP* sites for efficient Cre-mediated insertion into genomic DNA. *Genesis* **36,** 162–167 (2003).
24. Doyon, Y. *et al.* Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nature Methods* **8,** 74–79 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements State of Minnesota grant H001274506 to S.C.E. and D.F.V.; NIH GM63904 to S.C.E.; NIH grant P30DK084567 to S.C.E. and K.J.C.; NIH grant DK083219 to V.M.B.; Mayo Foundation; NIH DA032194 to K.J.C.; NIH grant R41HL108440 to D.F.C. and S.C.F.; NIH grant GM088424 to J.J.E.; NSF grant DBI0923827 to D.F.V.; General Research Fund (HKU771611, HKU771110, HKU769809M) from the Research Grant Council, The University of Hong Kong and the Tang King Yin Research Fund to A.C.M. and A.Y.H.L. We thank G. Davis for discussion on ssDNA use with custom restriction enzymes, H. J. Fadel for *in vitro* RNA synthesis, and S. Westcot for comments on this manuscript. We thank G. Moulder for help in DNA analyses and members of the Mayo Clinic Zebrafish Core Facility for excellent animal care.

Author Contributions W.T. designed and constructed full-length pT3TS-Tal vectors targeting *Danio rerio crhr1* and *crhr2* against exon sequences provided by K.J.C.; D.F.C. and S.C.F. designed and produced the pT3TS-Tal and GoldyTALEN cloning vectors. D.F.C. synthesized *crhr1* and *crhr2* TALEN mRNA from the pT3TS-Tal vectors. C.G.S. designed and assembled *ponzr1* TALENs and transferred TAL repeats from original *crhr1* and *crhr2* TALEN vectors into GoldyTALEN. T.L.P. and K.J.C. did initial *crhr1* TALEN microinjections. T.L.P. performed initial characterization of *crhr1* mutagenesis efficiency by PCR and restriction fragment length polymorphism analysis. T.L.P. and K.J.C. microinjected *crhr2* TALEN and *loxP* oligonucleotides. T.L.P. performed initial characterization by PCR demonstrating *loxP* integration. R.G.K. fully characterized efficiency and sequence of somatic *loxP* insertions in the *crhr2* locus. K.J.C. screened adult fin clips for *mloxP* integrations into *crhr2*. T.L.P. screened F1 offspring for *mloxP* integrations into *crhr2* and together with J.M.C. cloned and sequenced integration events. K.J.C. designed experiments associated for *crhr1* and *crhr2* modification. K.J.C. selected *loxP* mutant JTZ17 for integration. A.C.M. developed initial zebrafish genetic testing, TALEN cell-free assays, and ssDNA HDR protocols. S.G.P. conducted the cell-free TALEN endonuclease assay. A.Y.H.L. contributed to the design of initial TALEN and ssDNA HDR experiments. Y.W. and J.J.E. conducted biallelic conversion TALEN experiments in somatic and germline tests. D.F.V. and S.C.E. initiated the strategy to use custom restriction enzymes for genome editing in zebrafish. S.C.E. developed the plan for HDR targeting using ssDNAs, conducted overall project design and data analysis, and wrote the initial manuscript text. All authors contributed to manuscript composition. V.M.B. and J.M.C. conducted *ponzr1* and *crhr1* TALEN scaffold comparison experiments. V.M.B. and K.J.C. designed ssDNA oligonucleotides for HDR experiments. V.M.B. and J.M.C. injected and screened the *EcoRV* and *mloxP* HDR experiments at the *ponzr1* locus. J.M.C. ran quantitative data assessments and statistical analyses. V.M.B. and J.M.C. made first drafts and legends of Figs 1, 3 and 5. V.M.B. and J.M.C. conducted fin biopsy analyses of the *ponzr1* locus. V.M.B. completed the analysis of *ponzr1* germline transmission with assistance from J.M.C.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.C.E. (ekker.stephen@mayo.edu).

doi:10.1038/nature11582

Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease

A list of authors and their affiliations appears at the end of the paper.

Crohn's disease and ulcerative colitis, the two common forms of inflammatory bowel disease (IBD), affect over 2.5 million people of European ancestry, with rising prevalence in other populations[1]. Genome-wide association studies and subsequent meta-analyses of these two diseases[2,3] as separate phenotypes have implicated previously unsuspected mechanisms, such as autophagy[4], in their pathogenesis and showed that some IBD loci are shared with other inflammatory diseases[5]. Here we expand on the knowledge of relevant pathways by undertaking a meta-analysis of Crohn's disease and ulcerative colitis genome-wide association scans, followed by extensive validation of significant findings, with a combined total of more than 75,000 cases and controls. We identify 71 new associations, for a total of 163 IBD loci, that meet genome-wide significance thresholds. Most loci contribute to both phenotypes, and both directional (consistently favouring one allele over the course of human history) and balancing (favouring the retention of both alleles within populations) selection effects are evident. Many IBD loci are also implicated in other immune-mediated disorders, most notably with ankylosing spondylitis and psoriasis. We also observe considerable overlap between susceptibility loci for IBD and mycobacterial infection. Gene co-expression network analysis emphasizes this relationship, with pathways shared between host responses to mycobacteria and those predisposing to IBD.

We conducted an imputation-based association analysis using autosomal genotype-level data from 15 genome-wide association studies (GWAS) of Crohn's disease and/or ulcerative colitis (Supplementary Fig. 1 and Supplementary Table 1). We imputed 1.23 million single-nucleotide polymorphisms (SNPs) from the HapMap3 reference set (Supplementary Methods 1a), resulting in a high-quality data set with reduced genome-wide inflation (Supplementary Figs 2 and 3) compared with previous meta-analyses of subsets of these data[2,3]. The imputed GWAS data identified 25,075 SNPs that were associated ($P < 0.01$) with at least one of the Crohn's disease, ulcerative colitis, or combined IBD analyses. A meta-analysis of GWAS data with Immunochip[6] validation genotypes from an independent, newly genotyped set of 14,763 Crohn's disease cases, 10,920 ulcerative colitis cases and 15,977 controls was performed (Supplementary Fig. 1 and Supplementary Table 1). Principal-components analysis resolved geographic stratification, as well as Jewish and non-Jewish ancestry (Supplementary Fig. 4), and reduced inflation to a level consistent with residual polygenic risk, rather than other confounding effects (from a median test statistic inflation (λ_{GC}) = 2.00 to λ_{GC} = 1.23 when analysing all IBD samples; Supplementary Fig. 5 and Supplementary Methods 1b).

Our meta-analysis of the GWAS and Immunochip data identified 193 statistically independent signals of association at genome-wide significance ($P < 5 \times 10^{-8}$) in at least one of the three analyses (Crohn's disease, ulcerative colitis, IBD). Because some of these signals (Supplementary Fig. 6) probably represent associations to the same underlying functional unit, we merged these signals (Supplementary Methods 1b) into 163 regions, 71 of which are reported here for the first time (Table 1 and Supplementary Table 2). Fig. 1a shows the relative contributions of each locus to the total variance explained in ulcerative colitis and Crohn's disease. We have increased the total disease variance explained (variance being subject to fewer assumptions than heritability[7]) from 8.2% to 13.6% in Crohn's disease and from 4.1% to 7.5% in ulcerative colitis (Supplementary Methods 1c). Consistent with previous studies, our IBD risk loci seem to act independently, with no significant evidence of deviation from an additive combination of log odds ratios.

Our combined genome-wide analysis of Crohn's disease and ulcerative colitis enables a more comprehensive analysis of disease specificity than was previously possible. A model-selection analysis (Supplementary Methods 1c) showed that 110 out of 163 loci are associated with both disease phenotypes; 50 of these have an indistinguishable effect size in ulcerative colitis and Crohn's disease, whereas 60 show evidence of heterogeneous effects (Table 1). Of the remaining loci, 30 are classified as Crohn's-disease-specific and 23 as ulcerative-colitis-specific. However, 43 of these 53 loci show the same direction of effect in the non-associated disease (Fig. 1b; overall $P = 2.8 \times 10^{-6}$). Risk alleles at two Crohn's disease loci, *PTPN22* and *NOD2*, show significant ($P < 0.005$) protective effects in ulcerative colitis, exceptions that may reflect biological differences between the two diseases. This degree of sharing of genetic risk suggests that nearly all of the biological mechanisms involved in one disease have some role in the other.

The large number of IBD associations, far more than reported for any other complex disease, increases the power of network-based analyses to prioritize genes within loci. We investigated the IBD loci using functional annotation and empirical gene network tools (Supplementary Table 2). Compared with previous analyses that identified candidate genes in 35% of loci[2,3] our updated GRAIL[8] -connectivity network identifies candidates in 53% of loci, including increased statistical significance for 58 of the 73 candidates from previous analyses. The new candidates come not only from genes within newly identified loci, but also integrate additional genes from previously established loci (Fig. 1c). Only 29 IBD-associated SNPs are in strong linkage disequilibrium ($r^2 > 0.8$) with a missense variant in the 1000 Genomes Project data, which reinforces previous evidence that a large fraction of risk for complex disease is driven by non-coding variation. By contrast, 64 IBD-associated SNPs are in linkage disequilibrium with variants known to regulate gene expression (Supplementary Table 2). Overall, we highlighted a total of 300 candidate genes in 125 loci, of which 39 contained a single gene supported by two or more methods.

Seventy per cent (113 out of 163) of the IBD loci are shared with other complex diseases or traits, including 66 among the 154 loci previously associated with other immune-mediated diseases[9], which is 8.6-times the number that would be expected by chance ($P < 10^{-16}$; Fig. 2a and Supplementary Fig. 7). Such enrichment cannot be attributed to the immune-mediated focus of the Immunochip (Supplementary Methods 4 and Supplementary Fig. 8), as the analysis is based on our combined GWAS–Immunochip data. Comparing overlaps with specific diseases is confounded by the variable power in studies of different diseases. For instance, although type 1 diabetes shares the largest number of loci (20 out of 39; tenfold enrichment) with IBD, this is partially driven by the large number of known type 1 diabetes associations. Indeed, seven other immune-mediated diseases

Table 1 | Crohn's disease-specific, ulcerative colitis-specific and IBD general loci

Crohn's disease

Chr.	Position (Mb)	SNP	Key genes (+ no. of additional genes in locus)
1	78.62	rs17391694	(5)
1	114.3	**rs6679677**‖	***PTPN22*** ¶ (8)
1	120.45	rs3897478	*ADAM30* (5)
1	172.85	**rs9286879**	***FASLG***,***TNFSF18*** (0)
2	27.63	**rs1728918**	*UCN* (23)
2	62.55	rs10865331	(3)
2	231.09	**rs6716753**	*SP140* (5)
2	234.15	**rs12994997**	***ATG16L1*** ¶ (8)
4	48.36	rs6837335	(6)
4	102.86	rs13126505	(1)
5	55.43	rs10065637	***IL6ST***, ***IL31RA*** (1)
5	72.54	rs7702331	(4)
5	173.34	rs17695092	*CPEB4* (2)
6	21.42	rs12663356	(3)
6	31.27	**rs9264942**	(22)
6	127.45	rs9491697	(3)
6	128.24	**rs13204742**	(2)
6	159.49	**rs212388**	*TAGAP* (5)
7	26.88‡	rs10486483	(2)
7	28.17	rs864745	*CREB5*, *JAZF1* (1)
8	90.87	rs7015630	*RIPK2* (4)
8	129.56	**rs6651252**	0
13	44.45	**rs3764147**	*LACC1* (3)
15	38.89	rs16967103	***RASGRP1***, ***SPRED1*** (2)
16	50.66†	**rs2066847**‖	***NOD2*** ¶ (6)
17	25.84	**rs2945412**	***LGALS9***, ***NOS2*** (3)
19	1.12	**rs2024092**	*GPX4*, *HMHA1* (20)
19	46.85‡	rs4802307	(9)
19	49.2	**rs516246**	*FUT2*, (25)
21	34.77	**rs2284553**	***IFNGR2***, ***IFNAR1*** (10)

Ulcerative colitis

Chr.	Position (Mb)	SNP	Key genes (+ no. of additional genes in locus)
1	2.5	rs10797432	***TNFRSF14*** (10)
1	20.15†	**rs6426833**	(9)
1	200.09	**rs2816958**	(3)
2	198.65	rs1016883	*RFTN2*, *PLCL1* (7)
2	199.70*	**rs17229285**	0
3	53.05	rs9847710	*PRKCD*, *ITIH4* (8)
4	103.51	rs3774959	***NFKB1***, ***MANBA*** (2)
5	0.59	rs11739663	*SLC9A3* (8)
5	134.44	rs254560	(6)
6	32.595	**rs6927022**	(15)
7	2.78	**rs798502**	***CARD11***, ***GNA12*** (5)
7	27.22‡	rs4722672	(14)
7	107.45*	**rs4380874**	*DLD* (9)
7	128.57	**rs4728142**	***IRF5*** ¶ (13)
11	96.02	rs483905	*JRKL*, *MAML2* (2)
11	114.38	**rs561722**	*NXPE1*, *NXPE4* (5)
15	41.55	rs28374715	(11)
16	30.47	rs11150589	***ITGAL*** (20)
16	68.58	rs1728785	*ZFP90* (6)
17	70.64	rs7210086	(3)
19	47.12‡	rs1126510	*CALM3* (14)
20	33.8	rs6088765	(11)
20	43.06	**rs6017342**	*ADA*, *HNF4A* (9)

IBD

Chr.	Position (Mb)	SNP	Key genes (+ no. of additional genes in locus)
1	1.24	**rs12103**	*TNFRSF18*, *TNFRSF4* (30)
1	8.02	**rs35675666**	*TNFRSF9* (6)
1	22.7	**rs12568930**§	(3)
1	67.68†	**rs11209026**§	***IL23R*** ¶ (5)
1	70.99	rs2651244§	(3)
1	151.79	**rs4845604**§	*RORC* (14)
1	155.67	rs670523§	(31)
1	160.85	rs4656958§	***CD48*** (15)
1	161.47	**rs1801274**§	***FCGR2A***, ***FCGR2B*** & ***FCGR3A*** (13)
1	197.6	**rs2488389**	C1orf53 (2)
1	200.87	**rs7554511**	*KIF21B* (6)
1	206.93	**rs3024505**§	*IL10* (10)
2	25.12	**rs6545800**§	*ADCY3* (6)
2	28.61	**rs925255**§	*FOSL2*, *BRE* (1)
2	43.81	rs10495903§	(5)
2	61.2	**rs7608910**	***REL*** (9)
2	65.67	rs6740462	*SPRED2* (1)
2	102.86*	**rs917997**§	*IL18RAP*, *IL1R1* (7)
2	163.1	rs2111485	*IFIH1* (5)
2	191.92	rs1517352	***STAT1***, ***STAT4*** (2)
2	219.14	rs2382817	(15)
2	241.57*	**rs3749171**§	***GPR35*** (12)
3	18.76	**rs4256159**§	0
3	48.96†	**rs3197999**	***MST1***, ***PFKB4*** (63)
4	74.85	rs2472649§	(11)
4	123.22	**rs7657746**	***IL2***, ***IL21*** (2)
5	10.69	rs2930047	***DAP*** (2)
5	40.38†	**rs11742570**§	*PTGER4* (1)
5	96.24	**rs1363907**	*ERAP2*, *ERAP1* (3)
5	130.01	rs4836519§	(1)
5	131.19*	**rs2188962**§	*IBD5* locus (18)
5	141.51	**rs6863411**§	*SPRY4*, *NDFIP1* (5)
5	150.27	**rs11741861**§	***IRGM*** ¶ (10)
5	158.8†	**rs6871626**§	***IL12B*** (3)
5	176.79	rs12654812	***DOK3*** (17)
6	14.71	rs17119	0
6	20.77*	**rs9358372**§	(2)
6	90.96	rs1847472	(1)
6	106.43	**rs6568421**§	(2)
6	111.82	**rs3851228**	***TRAF3IP2*** (4)
6	138	**rs6920220**§	*TNFAIP3* (1)

IBD

Chr.	Position (Mb)	SNP	Key genes (+ no. of additional genes in locus)
10	35.3	**rs11010067**§	***CREM*** (3)
10	59.99	rs2790216	*CISD1*, *IPMK* (2)
10	64.51†	**rs10761659**§	(3)
10	75.67	rs2227564§	(13)
10	81.03	**rs1250546**§	(5)
10	82.25	**rs6586030**§	*TSPAN14*, C10orf58 (4)
10	94.43	rs7911264	(4)
10	101.28	**rs4409764**	*NKX2-3* (6)
11	1.87	rs907611	***TNNI2***, ***LSP1*** (17)
11	58.33	rs10896794	*CNTF*, *LPXN* (8)
11	60.77	**rs11230563**	***CD6*** (14)
11	61.56	**rs4246215**§	(15)
11	64.12	rs559928	***CCDC88B*** (23)
11	65.65	rs2231884§	*RELA* (25)
11	76.29	**rs2155219**§	(5)
11	87.12	rs6592362	(1)
11	118.74	rs630923§	*CXCR5* (17)
12	12.65	rs11612508§	*LOH12CR1* (8)
12	40.77*	**rs11564258**§	*MUC19* (1)
12	48.2	rs11168249§	*VDR* (8)
12	68.49	**rs7134599**§	***IFNG*** (3)
13	27.52	**rs17085007**§	(2)
13	40.86†	**rs941823**§	(3)
13	99.95	**rs9557195**	*GPR183*, *GPR18* (6)
14	69.27	rs194749§	*ZFP36L1* (4)
14	75.7	rs4899554§	***FOS***, ***MLH3*** (6)
14	88.47	**rs8005161**	***GPR65***, ***GALC*** (1)
15	67.43	**rs17293632**§	*SMAD3* (2)
15	91.17	rs7495132	*CRTC3* (3)
16	11.54*	**rs529866**§	***SOCS1***, ***LITAF*** (11)
16	23.86	rs7404095	***PRKCB*** (5)
16	28.6	**rs26528**§	*IL27* (14)
16	86	rs10521318§	*IRF8* (4)
17	32.59	**rs3091316**§	***CCL13***, ***CCL2*** (5)
17	37.91	**rs12946510**	*ORMDL3* (16)
17	40.53	**rs12942547**§	***STAT3*** (15)
17	57.96	**rs1292053**§	*TUBD1*, *RPS6KB1* (9)
18	12.8	**rs1893217**§	(6)
18	46.39	rs7240004§	*SMAD7* (2)
18	67.53	rs727088	***CD226*** (2)
19	10.49*	**rs11879191**	***TYK2*** (27)

Table 1 | Continued

		IBD	
Chr.	Position (Mb)	SNP	Key genes (+ no. of additional genes in locus)
6	143.9	rs12199775	*PHACTR2* (5)
6	167.37	**rs1819333**§	***CCR6***, ***RPS6KA2*** (4)
7	50.245*	**rs1456896**	*ZPBP*, *IKZF1* (4)
7	98.75	rs9297145	*SMURF1* (6)
7	100.34	**rs1734907**§	***EPO*** (21)
7	116.89	rs38904§	(6)
8	126.53	**rs921720**§	*TRIB1* (1)
8	130.62	rs1991866	(2)
9	4.98	**rs10758669**	***JAK2*** (4)
9	93.92	rs4743820§	*NFIL3* (2)
9	117.60†	**rs4246905**	*TNFSF15* (4)
9	139.32*	**rs10781499**§	***CARD9*** (22)
10	6.08	rs12722515§	***IL2RA***, ***IL15RA*** (6)
10	30.72	rs1042058§	***MAP3K8*** (3)
19	33.73	**rs17694108**	*CEBPG (8)*
19	55.38	rs11672983	(19)
20	30.75	rs6142618§	*HCK* (10)
20	31.37	rs4911259	*DNMT3B* (8)
20	44.74	**rs1569723**§	***CD40*** (13)
20	48.95	rs913678	*CEBPB* (5)
20	57.82	rs259964	*ZNF831*, *CTSZ* (5)
20	62.34	**rs6062504**	*TNFRSF6B* (26)
21	16.81	**rs2823286**§	0
21	40.46	**rs2836878**§	(3)
21	45.62	**rs7282490**	*ICOSLG* (9)
22	21.92	**rs2266959**	(13)
22	30.43	**rs2412970**	***LIF***, ***OSM*** (9)
22	39.69*	**rs2413583**§	*TAB1* (18)

The position given is the middle of the locus window, with all positions relative to human reference genome GRCh37. Bolded rs numbers indicate SNPs with P values less than 1×10^{-13}. Grey shading indicates newly discovered loci. Listed are genes implicated by one or more candidate gene approaches. Bolded genes have been implicated by two or more candidate gene approaches. For each locus, the top two candidate genes are listed. A complete listing of gene prioritization is provided in Supplementary Table 2. *Additional genome-wide significant associated SNP in the region. †Two or more additional genome-wide significant SNPs in the region. ‡These regions have overlapping but distinct ulcerative colitis and Crohn's disease signals. §Heterogeneity of odds ratios. ‖Crohn's disease risk allele is significantly protective in ulcerative colitis. ¶Gene for which functional studies of associated alleles have been reported. Chr., chromosome; Mb, megabase.

show stronger enrichment of overlap, with the largest being ankylosing spondylitis (8 out of 11; 13-fold) and psoriasis (14 out of 17; 14-fold).

IBD loci are also markedly enriched (4.9-fold; $P < 10^{-4}$) in genes involved in primary immunodeficiencies (PIDs; Fig. 2a), which are characterized by a dysfunctional immune system resulting in severe infections[10]. Genes implicated in this overlap correlate with reduced levels of circulating T cells (*ADA*, *CD40*, *TAP1*, *TAP2*, *NBN*, *BLM*, *DNMT3B*) or of specific subsets, such as T-helper cells producing IL-17 ($\mathrm{T_H}17$ cells) (*STAT3*), memory (*SP110*) or regulatory T cells (*STAT5B*). The subset of PID genes leading to Mendelian susceptibility to mycobacterial disease (MSMD)[10–12] is enriched still further; six of the eight known autosomal genes linked to MSMD are located within IBD loci (*IL12B*, *IFNGR2*, *STAT1*, *IRF8*, *TYK2*, *STAT3*; 46-fold enrichment; $P = 1.3 \times 10^{-6}$), and a seventh, *IFNGR1*, narrowly missed genome-wide significance ($P = 6 \times 10^{-8}$). Overlap with IBD is also seen in complex mycobacterial disease; we find IBD associations in seven out of eight loci identified by leprosy GWAS[13], including six cases in which the same SNP is implicated. Furthermore, genetic defects in *STAT3* (refs 14, 15) and *CARD9* (ref. 16), also within IBD loci, lead to PIDs involving skin infections with *Staphylococcus* and candidiasis, respectively. The comparative effects of IBD and infectious-disease-susceptibility-risk alleles on gene function and expression are summarized in Supplementary Table 3, and include both opposite (for example, *NOD2* and *STAT3*; Supplementary Fig. 9) and similar (for example, *IFNGR2*) directional effects.

To extend our understanding of the fundamental biology of IBD pathogenesis we conducted searches across the IBD locus list: (1) for enrichment of specific Gene Ontology terms and canonical pathways; (2) for evidence of selective pressure acting on specific variants and pathways; and (3) for enrichment of differentially expressed genes across immune-cell types. We tested the 300 prioritized genes (see above) for enrichment in Gene Ontology terms (Supplementary Methods 4a) and identified 286 Gene Ontology terms and 56 pathways demonstrating significant enrichment in genes contained within IBD loci (Supplementary Figs 10 and 11 and Supplementary Table 4). Excluding high-level Gene Ontology categories such as 'immune system processes' ($P = 3.5 \times 10^{-26}$), the most significantly enriched term is regulation of cytokine production ($P = 2.7 \times 10^{-24}$), specifically

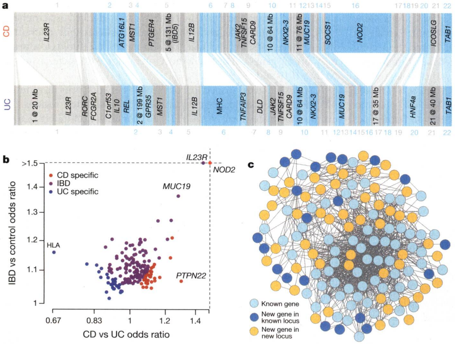


Figure 1 | The IBD genome. **a**, Variance explained by the 163 IBD loci. Each bar, ordered by genomic position, represents an independent locus. The width of the bar is proportional to the variance explained by that locus in Crohn's disease (CD) and ulcerative colitis (UC). Bars are connected together if they are identified as being associated with both phenotypes, and loci are labelled if they explain more than 1% of the total variance explained by all loci for that phenotype. Labels are either the best-supported candidate gene in Table 1, or the chromosome and position of the locus if either no, or multiple, well-supported candidates exist. **b**, The 193 independent signals, plotted by total IBD odds ratio and phenotype specificity (measured by the odds ratio of Crohn's disease relative to ulcerative colitis), and coloured by their IBD phenotype classification from Table 1. Note that many loci (for example, *IL23R*) show very different effects in Crohn's disease and ulcerative colitis despite being strongly associated to both. **c**, GRAIL network for all genes with GRAIL $P < 0.05$. Genes included in our previous GRAIL networks in both phenotypes are shown in light blue, newly connected genes in previously identified loci in dark blue, and genes from newly associated loci in gold. The gold genes reinforce the previous network (light blue) and expand it to include dark blue genes.

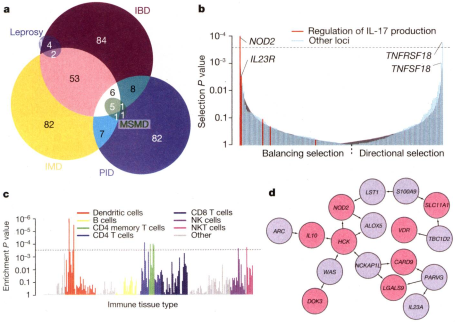


Figure 2 | Dissecting the biology of IBD. **a**, Number of overlapping IBD loci with other immune-mediated diseases (IMD), leprosy and Mendelian PIDs. Within PID, we highlight MSMD. **b**, Signals of selection at IBD SNPs, from strongest balancing on the left to strongest directional on the right. The grey curve shows the 95% confidence interval for randomly chosen frequency-matched SNPs, illustrating our overall enrichment ($P = 5.5 \times 10^{-6}$), and the dashed line represents the Bonferroni significance threshold. SNPs highlighted in red are annotated as being involved in the regulation of IL-17 production, a key IBD functional term related to bacterial defence, and are enriched for balancing selection. **c**, Evidence of enrichment in IBD loci of differentially expressed genes from various immune tissues. Each bar represents the empirical *P* value in a single tissue, and the colours represent different cell type groupings. The dashed line is Bonferroni-corrected significance for the number of tissues tested. **d**, *NOD2*-focused cluster of the IBD causal sub-network. Pink genes are in IBD-associated loci, blue are not. Arrows indicate inferred causal direction of regulation of expression.

interferon-γ, interleukin (IL)-12, tumour-necrosis factor-α and IL-10 signalling. Lymphocyte activation was the next most significant ($P = 1.8 \times 10^{-23}$), with activation of T cells, B cells and natural killer (NK) cells being the strongest contributors to this signal. Strong enrichment was also seen for response to molecules of bacterial origin ($P = 2.4 \times 10^{-20}$), and for the Kyoto Encyclopedia of Genes and Genomes (KEGG) JAK-STAT signalling pathway ($P = 4.8 \times 10^{-15}$). We note that no enriched terms or pathways showed specific evidence of Crohn's disease or ulcerative colitis specificity.

As infectious organisms are known to be among the strongest agents of natural selection, we investigated whether the IBD-associated variants are subject to selective pressures (Supplementary Table 5 and Supplementary Methods 4c). Directional selection would imply that the balance between these forces shifted in one direction over the course of human history, whereas balancing selection would suggest an allele-frequency-dependent scenario typified by host–microbe co-evolution, as can be observed with parasites. Two SNPs show Bonferroni-significant selection: the most significant signal, in *NOD2*, is under balancing selection ($P = 5.2 \times 10^{-5}$), and the second most significant, in the receptor *TNFRSF18*, showed directional selection ($P = 8.9 \times 10^{-5}$). The next most significant variants were in the ligand of that receptor, *TNFSF18* (directional; $P = 5.2 \times 10^{-4}$), and *IL23R* (balancing; $P = 1.5 \times 10^{-3}$). As a group, the IBD variants show significant enrichment in selection (Fig. 2b) of both types ($P = 5.5 \times 10^{-6}$). We discovered an enrichment of balancing selection (Fig. 2b) in genes annotated with the Gene Ontology term 'regulation of interleukin-17 production' ($P = 1.4 \times 10^{-4}$). The important role of IL-17 in both bacterial defence and autoimmunity suggests a key role for balancing selection in maintaining the genetic relationship between inflammation and infection, and this is reinforced by a nominal enrichment of balancing selection in loci annotated with the broader Gene Ontology term 'defense response to bacterium' ($P = 0.007$).

We tested for enrichment of cell-type expression specificity of genes in IBD loci in 223 distinct sets of sorted, mouse-derived immune cells from the Immunological Genome Consortium[17]. Dendritic cells showed the strongest enrichment, followed by weaker signals that support the Gene Ontology analysis, including CD4$^+$ T cells, NK cells and NKT cells (Fig. 2c). Notably, several of these cell types express genes near our IBD associations much more specifically when stimulated; our strongest signal, a lung-derived dendritic cell, had $P_{\text{stimulated}} < 1 \times 10^{-6}$ compared with $P_{\text{unstimulated}} = 0.0015$, consistent with an important role for cell activation.

To further our goal of identifying likely causal genes within our susceptibility loci and to elucidate networks underlying IBD pathogenesis, we screened the associated genes against 211 co-expression modules identified from weighted gene co-expression network analyses[18], conducted with large gene-expression data sets from multiple tissues[19–21]. The most significantly enriched module comprised 523 genes from omental adipose tissue collected from morbidly obese patients[19], which was found to be 2.9-fold enriched for genes in the IBD-associated loci ($P = 1.1 \times 10^{-13}$; Supplementary Fig. 12 and Supplementary Table 6). We constructed a probabilistic causal gene network using an integrative Bayesian network-reconstruction algorithm[22–24], which combines expression and genotype data to infer the direction of causality between genes with correlated expression. The intersection of this network and the genes in the IBD-enriched module defined a sub-network of genes enriched in bone marrow-derived macrophages ($P < 10^{-16}$) and is suggestive of dynamic interactions relevant to IBD pathogenesis. In particular, this sub-network featured close proximity among genes connected to host interaction with bacteria, notably *NOD2*, *IL10* and *CARD9*.

A *NOD2*-focused inspection of the sub-network prioritizes multiple additional candidate genes within IBD-associated regions. For example, a cluster near *NOD2* (Fig. 2d) contains multiple IBD genes implicated in the *Mycobacterium tuberculosis* response, including *SLC11A1*, *VDR* and *LGALS9*. Furthermore, both *SLC11A1* (also known as *NRAMP1*) and *VDR* have been associated with *M. tuberculosis* infection by candidate gene studies[25,26], and *LGALS9* modulates mycobacteriosis[27]. Of interest, *HCK* (located in our new locus on chromosome 20 at 30.75 megabases) is predicted to upregulate expression of both *NOD2* and *IL10*, an anti-inflammatory cytokine associated with Mendelian[28] and non-Mendelian [29] IBD. *HCK* has been linked to alternative, anti-inflammatory activation of monocytes (M2-group macrophages)[30]; although not identified in our aforementioned analyses, these data implicate *HCK* as the causal gene in this new IBD locus.

We report one of the largest genetic experiments involving a complex disease undertaken to date. This has increased the number of confirmed IBD susceptibility loci to 163, most of which are associated with both Crohn's disease and ulcerative colitis, and is substantially more than reported for any other complex disease. Even this large number of loci explains only a minority of the variance in disease risk, which suggests that other factors—such as rarer genetic variation not captured by GWAS or environmental exposures—make substantial contributions to pathogenesis. Most of the evidence relating to possible causal genes points to an essential role for host defence against infection in IBD. In this regard, the current results focus ever-closer attention on the interaction between the host mucosal immune system and microbes, both at the epithelial cell surface and within the gut lumen. In particular, they raise the question, in the context of this burden of IBD-susceptibility genes, of what triggers components of the commensal microbiota to switch from a symbiotic to a pathogenic relationship with the host. Collectively, our findings begin to shed light on these questions and provide a rich source of clues to the pathogenic mechanisms underlying this archetypal complex disease.

METHODS SUMMARY

We conducted a meta-analysis of GWAS data sets after imputation to the HapMap3 reference set, and aimed to replicate in the Immunochip data any SNPs with $P < 0.01$. We compared likelihoods of different disease models to assess whether each locus was associated with Crohn's disease, ulcerative colitis, or both. We used databases of expression quantitative trait loci SNPs and coding SNPs in linkage disequilibrium with our hit SNPs, as well as the network tools GRAIL and DAPPLE, and a co-expression network analysis to prioritize candidate genes in our loci. Gene Ontology, the Immunological Genome Project (ImmGen) mouse immune-cell-expression resource, the TreeMix selection software and a Bayesian causal network analysis were used to functionally annotate these genes.

Received 17 May; accepted 12 September 2012.

1. Molodecky, N. A. *et al.* Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* **142,** 46–54 (2012).
2. Anderson, C. A. *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nature Genet.* **43,** 246–252 (2011).
3. Franke, A. *et al.* Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nature Genet.* **42,** 1118–1125 (2010).
4. Khor, B., Gardet, A. & Xavier, R. J. Genetics and pathogenesis of inflammatory bowel disease. *Nature* **474,** 307–317 (2011).
5. Cho, J. H. & Gregersen, P. K. Genomics and the multifactorial nature of human autoimmune disease. *N. Engl. J. Med.* **365,** 1612–1623 (2011).
6. Cortes, A. & Brown, M. A. Promise and pitfalls of the Immunochip. *Arthritis Res. Ther.* **13,** 101 (2011).
7. Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability: genetic interactions create phantom heritability. *Proc. Natl Acad. Sci. USA* **109,** 1193–1198 (2012).
8. Raychaudhuri, S. *et al.* Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet.* **5,** e1000534 (2009).
9. Hindorff, L. A. *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl Acad. Sci. USA* **106,** 9362–9367 (2009).
10. Notarangelo, L. D. *et al.* Primary immunodeficiencies: 2009 update. *J. Allergy Clin. Immunol.* **124,** 1161–1178 (2009).
11. Bustamante, J., Picard, C., Boisson-Dupuis, S., Abel, L. & Casanova, J. L. Genetic lessons learned from X-linked Mendelian susceptibility to mycobacterial diseases. *Ann. NY Acad. Sci.* **1246,** 92–101 (2011).
12. Patel, S. Y., Doffinger, R., Barcenas-Morales, G. & Kumararatne, D. S. Genetically determined susceptibility to mycobacterial infection. *J. Clin. Pathol.* **61,** 1006–1012 (2008).
13. Zhang, F. *et al.* Identification of two new loci at *IL23R* and *RAB32* that influence susceptibility to leprosy. *Nature Genet.* **43,** 1247–1251 (2011).
14. Holland, S. M. *et al.* *STAT3* mutations in the hyper-IgE syndrome. *N. Engl. J. Med.* **357,** 1608–1619 (2007).
15. Minegishi, Y. *et al.* Dominant-negative mutations in the DNA-binding domain of STAT3 cause hyper-IgE syndrome. *Nature* **448,** 1058–1062 (2007).
16. Glocker, E. O. *et al.* A homozygous *CARD9* mutation in a family with susceptibility to fungal infections. *N. Engl. J. Med.* **361,** 1727–1735 (2009).
17. Hu, X. *et al.* Integrating autoimmune risk loci with gene-expression data identifies specific pathogenic immune cell subsets. *Am. J. Hum. Genet.* **89,** 496–506 (2011).
18. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4,** Article 17 (2005).
19. Greenawalt, D. M. *et al.* A survey of the genetics of stomach, liver, and adipose gene expression from a morbidly obese cohort. *Genome Res.* **21,** 1008–1016 (2011).
20. Emilsson, V. *et al.* Genetics of gene expression and its effect on disease. *Nature* **452,** 423–428 (2008).
21. Schadt, E. E. *et al.* Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* **6,** e107 (2008).
22. Chen, Y. *et al.* Variations in DNA elucidate molecular networks that cause disease. *Nature* **452,** 429–435 (2008).
23. Zhong, H. *et al.* Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. *PLoS Genet.* **6,** e1000932 (2010).
24. Zhu, J. *et al.* Increasing the power to detect causal associations by combining genotypic and expression data in segregating populations. *PLOS Comput. Biol.* **3,** e69 (2007).
25. Lewis, S. J., Baker, I. & Davey Smith, G. Meta-analysis of vitamin D receptor polymorphisms and pulmonary tuberculosis risk. *Int. J. Tuberc. Lung Dis.* **9,** 1174–1177 (2005).
26. Li, X. *et al.* *SLC11A1* (NRAMP1) polymorphisms and tuberculosis susceptibility: updated systematic review and meta-analysis. *PLoS ONE* **6,** e15831 (2011).
27. Kumar, D. *et al.* Genome-wide analysis of the host intracellular network that regulates survival of Mycobacterium tuberculosis. *Cell* **140,** 731–743 (2010).
28. Glocker, E. O. *et al.* Infant colitis–it's in the genes. *Lancet* **376,** 1272 (2010).
29. Franke, A. *et al.* Sequence variants in *IL10*, *ARPC2* and multiple other loci contribute to ulcerative colitis susceptibility. *Nature Genet.* **40,** 1319–1323 (2008).
30. Bhattacharjee, A., Pal, S., Feldman, G. M. & Cathcart, M. K. Hck is a key regulator of gene expression in alternatively activated human monocytes. *J. Biol. Chem.* **286,** 36709–36723 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank all the subjects who contributed samples and the physicians and nursing staff who helped with recruitment globally. UK case collections were supported by the National Association for Colitis and Crohn's disease; Wellcome Trust grant 098051 (L.J., C.A.A., J.C.B.); Medical Research Council UK; the Catherine McEwan Foundation; an NHS Research Scotland career fellowship (R.K.R.); Peninsula College of Medicine and Dentistry, Exeter; the National Institute for Health Research, through the Comprehensive Local Research Network, and through Biomedical Research Centre awards to Guy's & Saint Thomas' National Health Service Trust, King's College London, Addenbrooke's Hospital, University of Cambridge School of Clinical Medicine and to the University of Manchester and Central Manchester Foundation Trust. The British 1958 Birth Cohort DNA collection was funded by Medical Research Council grant G0000934 and Wellcome Trust grant 068545/Z/02, and the UK National Blood Service controls by the Wellcome Trust. The Wellcome Trust Case Control Consortium projects were supported by Wellcome Trust grants 083948/Z/07/Z, 085475/B/08/Z and 085475/Z/08/Z. North American collections and data processing were supported by funds to the National Institute of Diabetes, Digestive and Kidney diseases (NIDDK) IBD Genetics Consortium, which is funded by the following grants: DK062431 (S.R.B.), DK062422 (J.H.C.), DK062420 (R.H.D.), DK062432 (J.D.R.), DK062423 (M.S.S.), DK062413 (D.P.M.), DK076984 (M.J.D.), DK084554 (M.J.D. and D.P.M.) and DK062429 (J.H.C.). Additional funds were provided by funding to J.H.C. (DK062429-S1 and Crohn's & Colitis Foundation of America, Senior Investigator Award (5-2229)) and R.H.D. (CA141743). K.Y.H. is supported by the National Institutes of Health (NIH) MSTP TG T32GM07205 training award. Cedars-Sinai is supported by USPHS grant PO1DK046763 and the Cedars-Sinai F. Widjaja Inflammatory Bowel and Immunobiology Research Institute Research Funds, National Center for Research Resources (NCRR) grant M01-RR00425, UCLA/Cedars-Sinai/Harbor/Drew Clinical and Translational Science Institute (CTSI) Grant (UL1 TR000124-01), the Southern California Diabetes and Endocrinology Research Grant (DERC) (DK063491), The Helmsley Foundation (D.P.M.) and the Crohn's and Colitis Foundation of America (D.P.M.). R.J.X. and A.N.A. are funded by DK83756, AI062773, DK043351 and the Helmsley Foundation. The Netherlands Organization for Scientific Research supported R.K.W. with a clinical fellowship grant (90.700.281) and C.W. (VICI grant 918.66.620). C.W. is also supported by the Celiac Disease Consortium (BSIK03009). This study was also supported by the German Ministry of Education and Research through the National Genome Research Network, the Popgen biobank, through the Deutsche Forschungsgemeinschaft (DFG) cluster of excellence 'Inflammation at Interfaces' and DFG grant no. FR 2821/2-1. S.B. was supported by DFG BR 1912/6-1 and the Else Kröner-Fresenius-Stiftung (Else Kröner-Exzellenzstipendium 2010_EKES.32). Italian case collections were supported by the Italian Group for IBD and the Italian Society for Paediatric Gastroenterology, Hepatology and Nutrition and funded by the Italian Ministry of Health GR-2008-1144485. Activities in Sweden were supported by the Swedish Society of Medicine, Ihre Foundation, Örebro University Hospital Research Foundation, Karolinska Institutet, the Swedish National Program for IBD Genetics, the Swedish Organization for IBD, and the Swedish Medical Research Council. D.F. and S.V. are senior clinical investigators for the Funds for Scientific Research (FWO/FNRS) Belgium. We acknowledge a grant from Viborg Regional Hospital, Denmark. V. Andersen was supported by SHS Aabenraa, Denmark. We acknowledge funding provided by the Royal Brisbane and Women's Hospital Foundation, National Health and Medical Research Council, Australia and by the European Community (5th PCRDT). We acknowledge the following groups that provided biological samples or data for this study: the Inflammatory Bowel in South Eastern Norway (IBSEN) study group, the Norwegian Bone Marrow Donor Registry (NMBDR), the Avon Longitudinal Study of Parents and Children, the Human Biological Data Interchange and Diabetes UK, and Banco Nacional de ADN, Salamanca. This research also uses resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the NIDDK, National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation (JDRF) and supported by U01 DK062418. The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. KORA research was supported within the Munich Center of

Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

Author Contributions R.K.W., R.H.D., D.P.M., C.G.M., J.D.R., E.E.S., M.J.D., A.F., M.P. and S.V. contributed equally to the manuscript. J.H.C., J.C.B., R.K.W., R.H.D., D.P.M., A.F., M.P., C.G.M., J.D.R., S.V., M.D.A. and V. Annese conceived, designed and managed the study and managed the funding. J.H.C., J.C.B., L.J., S. Ripke, R.K.W., R.H.D., D.P.M., M.J.D., M.P. and C.G.M. were involved in manuscript preparation. J.H.C., J.C.B., L.J., S. Ripke, R.K.W., K.Y.H., C.A.A., J.E., K.N., S.L.S., S. Raychaudhuri, Z.W., C.A., A.C., G.B., M.H., X.H., B.Z., C.K.Z., H.Z., J.D.R., E.E.S. and M.J.D. performed or supervised statistical and computational analyses. R.K.W., R.H.D., D.P.M., J.C.L., L.P.S., Y.S., P.G., J.-P.A., T.A., L.A., A.N.A., V. Andersen, J.M.A., L.B., P.A.B., A.B., S.B., C.B., S.C., M.D.A., D.D.J., K.L.D., M.D., C.E., L.R.F., D.F., M.G., C.G., R.G., J.G., A.H., C.H., T.H.K., L.K., S.K., A.L., D.L., E.L., I.C.L., C.W.L., A.R.M., C.M., G.M., J.M., W.N., O.P., C.Y.P., U.P., N.J.P., M.R., J.I.R., R.K.R., J.D.S., M.S., J. Satsangi, S.S., L.A.S., J. Sventoraityte, S.R.T., M.T., H.W.V., M.D.V., C.W., D.C.W., J.W., R.J.X., S.Z., M.S.S., V. Annese, H.H., IIBDGC, S.R.B., J.D.R., G.R.S., C.G.M., A.F., M.P., S.V. and J.H.C. were involved in study subject recruitment and assembling phenotypic data. R.K.W., R.H.D., D.P.M., L.P.S., Y.S., M.M., I.C., E.T., T.B., D.E., K.F., T.H., K.D.T., C.G.M., A.F., M.P. and J.H.C. established DNA collections, genotyping and data management. All authors read and approved the final manuscript before submission.

Author Information Data have been deposited in the NCBI database of Genotypes and Phenotypes under accession numbers phs000130.v1.p1 and phs000345.v1.p1. Summary statistics for imputed GWAS are available at http://www.broadinstitute.org/mpg/ricopili/. Summary statistics for the meta-analysis markers are available at http://www.ibdgenetics.org/. The 523 causal gene network cytoscape file is available on request. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.H.C. (judy.cho@yale.edu).

Luke Jostins[1]*, Stephan Ripke[2,3]*, Rinse K. Weersma[4], Richard H. Duerr[5,6], Dermot P. McGovern[7,8], Ken Y. Hui[9], James C. Lee[10], L. Philip Schumm[11], Yashoda Sharma[12], Carl A. Anderson[1], Jonah Essers[13], Mitja Mitrovic[14,15], Kaida Ning[12], Isabelle Cleynen[16], Emilie Theatre[17,18], Sarah L. Spain[19], Soumya Raychaudhuri[20,21,22], Philippe Goyette[23], Zhi Wei[24], Clara Abraham[12], Jean-Paul Achkar[25,26], Tariq Ahmad[27], Leila Amininejad[28], Ashwin N. Ananthakrishnan[29], Vibeke Andersen[30], Jane M. Andrews[31], Leonard Baidoo[5], Tobias Balschun[32], Peter A. Bampton[33], Alain Bitton[34], Gabrielle Boucher[23], Stephan Brand[35], Carsten Büning[36], Ariella Cohain[37], Sven Cichon[38], Mauro D'Amato[39], Dirk De Jong[4], Kathy L. Devaney[29], Marla Dubinsky[40], Cathryn Edwards[41], David Ellinghaus[32], Lynnette R. Ferguson[42], Denis Franchimont[28], Karin Fransen[5,43], Richard Gearry[44,45], Michel Georges[17], Christian Gieger[46], Jürgen Glas[34], Talin Haritunians[8], Ailsa Hart[47], Chris Hawkey[48], Matija Hedl[12], Xinli Hu[20], Tom H. Karlsen[49], Limas Kupcinskas[50], Subra Kugathasan[51], Anna Latiano[52], Debby Laukens[53], Ian C. Lawrance[54], Charlie W. Lees[55], Edouard Louis[18], Gillian Mahy[56], John Mansfield[57], Angharad R. Morgan[42], Craig Mowat[58], William Newman[59], Orazio Palmieri[52], Cyriel Y. Ponsioen[60], Uros Potocnik[14,61], Natalie J. Prescott[19], Miguel Regueiro[5], Jerome I. Rotter[8], Richard K. Russell[62], Jeremy D. Sanderson[63], Miquel Sans[64,65], Jack Satsangi[55], Stefan Schreiber[32,66], Lisa A. Simms[67], Jurgita Sventoraityte[50], Stephan R. Targan[7], Kent D. Taylor[7,8], Mark Tremelling[68], Hein W. Verspaget[69], Martine De Vos[53], Cisca Wijmenga[43], David C. Wilson[62,70], Juliane Winkelmann[71], Ramnik J. Xavier[29,72], Sebastian Zeissig[66], Bin Zhang[37], Clarence K. Zhang[73], Hongyu Zhao[73], The International IBD Genetics Consortium (IIBDGC)†, Mark S. Silverberg[74], Vito Annese[52,75], Hakon Hakonarson[76,77], Steven R. Brant[78], Graham Radford-Smith[67,79], Christopher G. Mathew[19], John D. Rioux[23], Eric E. Schadt[37], Mark J. Daly[2,3], Andre Franke[32], Miles Parkes[10], Severine Vermeire[16,80], Jeffrey C. Barrett[1]* & Judy H Cho[9,12]*

[1]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1HH, UK. [2]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. [3]Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. [4]Department of Gastroenterology and Hepatology, University of Groningen and University Medical Center Groningen, Groningen 9700 RB, The Netherlands. [5]Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261, USA. [6]Department of Human Genetics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania 15261, USA. [7]F. Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute, Los Angeles, California 90048, USA. [8]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California 90048, USA. [9]Department of Genetics, Yale School of Medicine, New Haven, Connecticut 06520, USA. [10]Inflammatory Bowel Disease Research Group, Addenbrooke's Hospital, University of Cambridge, Cambridge CB2 0QQ, UK. [11]Department of Health Studies, University of Chicago, Chicago, Illinois 60637, USA. [12]Department of Internal Medicine, Section of Digestive Diseases, Yale School of Medicine, New Haven, Connecticut 06520, USA. [13]Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. [14]University of Maribor, Faculty of Medicine, Center for Human Molecular Genetics and Pharmacogenomics, Maribor 2000, Slovenia. [15]University Medical Center Groningen, Department of Genetics, Groningen 9700 RB, The Netherlands. [16]Department of Clinical and Experimental Medicine, Gastroenterology section, KU Leuven, Leuven 3000, Belgium. [17]Unit of Animal Genomics, Groupe Interdisciplinaire de Génoprotéomique Appliquée (GIGA-R) and Faculty of Veterinary Medicine, University of Liège, Liège 4000, Belgium. [18]Division of Gastroenterology, Centre Hospitalier Universitaire, Université de Liège, Liège 4000, Belgium. [19]Department of Medical and Molecular Genetics, Division of Genetics and Molecular Medicine, King's College London School of Medicine, Guy's Hospital, London SE1 9RT, UK. [20]Division of Rheumatology Immunology and Allergy, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [21]Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts 02142, USA. [22]Division of Genetics, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [23]Université de Montréal and the Montreal Heart Institute, Research Center, Montréal, Québec H1T 1C8, Canada. [24]Department of Computer Science, New Jersey Institute of Technology, Newark, New Jersey 07102, USA. [25]Department of Gastroenterology & Hepatology, Digestive Disease Institute, Cleveland Clinic, Cleveland, Ohio 44195, USA. [26]Department of Pathobiology, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio 44195, USA. [27]Peninsula College of Medicine and Dentistry, Exeter EX1 2LU, UK. [28]Erasmus Hospital, Free University of Brussels, Department of Gastroenterology, Brussels, 1070 Belgium. [29]Massachusetts General Hospital, Harvard Medical School, Gastroenterology Unit, Boston, Massachusetts 02114, USA. [30]Viborg Regional Hospital, Medical Department, Viborg 8800, Denmark. [31]Inflammatory Bowel Disease Service, Department of Gastroenterology and Hepatology, Royal Adelaide Hospital, and School of Medicine, University of Adelaide, Adelaide 5000, Australia. [32]Institute of Clinical Chemistry, Christian-Albrechts-University, Kiel 24105, Germany. [33]Department of Gastroenterology and Hepatology, Flinders Medical Centre and School of Medicine, Flinders University, Adelaide 5000, Australia. [34]Division of Gastroenterology, McGill University Health Centre, Royal Victoria Hospital, Montréal, Québec H3A 1A1, Canada. [35]Department of Medicine II, University Hospital Munich-Grosshadern, Ludwig-Maximilians-University, Munich 80336, Germany. [36]Department of Gastroenterology, Charité, Campus Mitte, Universitätsmedizin Berlin, Berlin 10117, Germany. [37]Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York City, New York 10029, USA. [38]Department of Genomics, Life & Brain Center, University Hospital Bonn, Bonn 53012, Germany. [39]Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm 14 183, Sweden. [40]Department of Pediatrics, Cedars Sinai Medical Center, Los Angeles, California 90048, USA. [41]Torbay Hospital, Department of Gastroenterology, Torbay, Devon TQ2 7AA, UK. [42]School of Medical Sciences, Faculty of Medical & Health Sciences, The University of Auckland, Auckland 1142, New Zealand. [43]University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen T9700 RB, The Netherlands. [44]Department of Medicine, University of Otago, Christchurch 8140, New Zealand. [45]Department of Gastroenterology, Christchurch Hospital, Christchurch 8011, New Zealand. [46]Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg 85764, Germany. [47]St Mark's Hospital, Watford Road, Harrow, Middlesex HA1 3UJ, UK. [48]Nottingham Digestive Diseases Centre, Queens Medical Centre, Nottingham NG7 1AW, UK. [49]Research Institute of Internal Medicine, Oslo University Hospital Rikshospitalet, Oslo 0424, Norway. [50]Kaunas University of Medicine, Department of Gastroenterology, Kaunas 44307, Lithuania. [51]Department of Pediatrics, Emory University School of Medicine, Atlanta, Georgia 30322, USA. [52]Unit of Gastroenterology, Istituto di Ricovero e Cura a Carattere Scientifico-Casa Sollievo della Sofferenza (IRCCS-CSS) Hospital, San Giovanni Rotondo 71013, Italy. [53]Ghent University Hospital, Department of Gastroenterology and Hepatology, Ghent 9000, Belgium. [54]School of Medicine and Pharmacology, The University of Western Australia, Fremantle, Western Australia 6009, Australia. [55]Gastrointestinal Unit, Molecular Medicine Centre, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK. [56]Department of Gastroenterology, The Townsville Hospital, Townsville, Queensland 4810, Australia. [57]Institute of Human Genetics, Newcastle University, Newcastle upon Tyne NE1 7RU, UK. [58]Department of Medicine, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK. [59]Genetic Medicine, MAHSC, University of Manchester, Manchester M13 9PL, UK. [60]Academic Medical Center, Department of Gastroenterology, Amsterdam 1105 AZ, The Netherlands. [61]University of Maribor, Faculty for Chemistry and Chemical Engineering, Maribor 2000, Slovenia. [62]Royal Hospital for Sick Children, Paediatric Gastroenterology and Nutrition, Glasgow G3 8SJ, UK. [63]Guy's & St Thomas' NHS Foundation Trust, St Thomas' Hospital, Department of Gastroenterology, London SE1 7EH, UK. [64]Department of Gastroenterology, Hospital Clinic/Institut d'Investigacions Biomédiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain. [65]Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBER-EHD), Barcelona 08036, Spain. [66]Department for General Internal Medicine, Christian-Albrechts-University, Kiel, Kiel 24118, Germany. [67]Inflammatory Bowel Diseases, Genetics and Computational Biology, Queensland Institute of Medical Research, Brisbane 4029, Australia. [68]Norfolk and Norwich University Hospital, Norwich NR4 7UY, UK. [69]Department of Gastroenterology, Leiden University Medical Center, Leiden 2333 ZA, The Netherlands. [70]Child Life and Health, University of Edinburgh, Edinburgh, Scotland EH9 1UW, UK. [71]Institute of Human Genetics and Department of Neurology, Technische Universität München, Munich 80336, Germany. [72]Center for Computational and Integrative Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [73]Department of Biostatistics, School of Public Health, Yale University, New Haven, Connecticut 06520, USA. [74]Mount Sinai Hospital Inflammatory Bowel Disease Centre, University of Toronto, Toronto, Ontario M5G 1X5, Canada. [75]Azienda Ospedaliero Universitaria (AOU) Careggi, Unit of Gastroenterology SOD2, Florence 50134, Italy. [76]Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA. [77]Department of Pediatrics, Center for Pediatric Inflammatory Bowel Disease, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA. [78]Meyerhoff Inflammatory Bowel Disease Center, Department of Medicine, School of Medicine, and Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland 21205, USA. [79]Department of Gastroenterology, Royal Brisbane and Women's Hospital, and School of Medicine, University of Queensland, Brisbane 4029, Australia. [80]Department of Gastroenterology, University Hospital Leuven, Leuven 3000, Belgium.

*These authors contributed equally to this work.
†Lists of participants and their affiliations appear in the Supplementary Information.

LETTER

doi:10.1038/nature11433

Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11

Manqing Li[1], Elaine Kao[1], Xia Gao[1], Hilary Sandig[1], Kirsten Limmer[1], Mariana Pavon-Eternod[2], Thomas E. Jones[2], Sebastien Landry[3], Tao Pan[2], Matthew D. Weitzman[3] & Michael David[1,4]

In mammals, one of the most pronounced consequences of viral infection is the induction of type I interferons, cytokines with potent antiviral activity. Schlafen (*Slfn*) genes are a subset of interferon-stimulated early response genes (ISGs) that are also induced directly by pathogens via the interferon regulatory factor 3 (IRF3) pathway[1]. However, many ISGs are of unknown or incompletely understood function. Here we show that human SLFN11 potently and specifically abrogates the production of retroviruses such as human immunodeficiency virus 1 (HIV-1). Our study revealed that SLFN11 has no effect on the early steps of the retroviral infection cycle, including reverse transcription, integration and transcription. Rather, SLFN11 acts at the late stage of virus production by selectively inhibiting the expression of viral proteins in a codon-usage-dependent manner. We further find that SLFN11 binds transfer RNA, and counteracts changes in the tRNA pool elicited by the presence of HIV. Our studies identified a novel antiviral mechanism within the innate immune response, in which SLFN11 selectively inhibits viral protein synthesis in HIV-infected cells by means of codon-bias discrimination.

SLFN genes encode a family of proteins limited to mammalian organisms. Nine murine and six human *SLFN* genes share a conserved NH2-terminus containing a putative AAA-domain, and long *SLFN* genes possess motifs resembling DNA/RNA helicase domains, a trait they share with the nucleic acid sensors RIG-I and MDA-5[2]. Beyond that, SLFN proteins harbour no sequence similarity to other proteins. *In vivo*, short and long murine SLFN proteins inhibit T-cell development[3–5], and levels of murine SLFN proteins are elevated after infection with *Brucella* or *Listeria*[4]. Lipopolysaccharide, poly-inosine-cytosine (poly-IC) or interferon (IFN)-α/β treatment of macrophages results in induction of several murine *Slfn* genes (our unpublished results). Treatment of human foreskin fibroblasts with IFN-β, poly-IC or poly-dAdT revealed similar induction of *SLFN* genes (Supplementary Fig. 1a), and human SLFN5 and SLFN11 were consistently the most prominent family members (Supplementary Fig. 1b). Notably, we observed a striking difference in *SLFN* levels between HEK293 (293) and HEK293T (293T) cells (Supplementary Fig. 1b), and exploited this differential expression to focus on SLFN11 for further studies. We further used *SLFN11*-targeted short hairpin RNA to generate stable 293 cells that specifically lack SLFN11 expression (293shRNA$^{\text{SLFN}}$) (Supplementary Fig. 1c, d).

To test whether lack of SLFN11 in 293shRNA$^{\text{SLFN}}$ or 293T cells alters their ability to subdue viral infections, we infected these cells with vesicular stomatitis virus (VSV)-G pseudotyped HIV (HIV$^{\text{VSV-G}}$), or amphotropic murine stem cell virus (MSCV), adeno-associated virus (AAV), or herpes simplex virus (HSV). HIV$^{\text{VSV-G}}$-infected cells expressed luciferase after integration of the viral complementary DNAs into the host genome. Regardless of SLFN11 expression, all cell lines had comparable luciferase levels after HIV$^{\text{VSV-G}}$ infection (Supplementary Fig. 2). A similar lack of influence of SLFN11 was observed when cells were infected with MSCV, AAV or HSV (not shown).

HEK293T cells are used as packaging cells for production of retroviruses, and we therefore considered the possibility that virus production rather than the response towards them is afflicted by SLFN11. Indeed, 293T cells produced markedly higher HIV$^{\text{VSV-G}}$ (Fig. 1b) or MCSV (Supplementary Fig. 3a) titres than 293 cells from the viral vectors pNL4-3.Luc.R^{+}E^{-} or MSCV-IRES-GFP, respectively. Most importantly, this increase in viral titre was also clearly evident in 293shRNA$^{\text{SLFN}}$ cells, whereas 293 and 293shRNA$^{\text{Ctl}}$ cells produced the same low levels of virus (Fig. 1b and Supplementary Fig. 3a). Notably, the modulation of virus production is limited to particular viruses, as fabrication of retroviruses (Fig. 1b and Supplementary Fig. 3a), but not of AAV (Supplementary Fig. 3c), was affected by SLFN11. We also did not observe any modulation of ISGs such as ISG15, ISG54 or APOBEC3G (not shown) as a consequence of SLFN11 expression, supporting the notion that SLFN11 does not create a general virus-resistant phenotype.

To corroborate that the observed differences are attributable to dissimilar SLFN11 expression, we expressed full-length SLFN11 (amino acids 1–901) in 293T cells and analysed their ability to produce HIV$^{\text{VSV-G}}$. Indeed, SLFN11 strongly inhibited HIV$^{\text{VSV-G}}$ (Fig. 1a) or MSCV (Supplementary Fig. 3b) production from 293T cells, with the inhibitory activity residing in the AAA-domain-containing, amino-terminal region (SLFN11-N; amino acids 1–579). No effect of the isolated carboxy-terminal region (SLFN11-C; amino acids 523–901) harbouring the putative helicase sequence was observed (Fig. 1a and Supplementary Fig. 3b). Intriguingly, SLFN5 failed to inhibit retrovirus production but yielded slightly elevated viral titres, illustrating specificity among SLFN proteins in their antiviral activity.

To discern whether SLFN11 reduced the number or the viability of released virus, we measured p24 capsid and viral RNA (vRNA) levels in supernatants of pNL4-3.Luc.R^{+}E^{-}-transfected, HIV$^{\text{VSV-G}}$-producing cells. Extracellular p24 (Fig. 1c and d) or vRNA (Fig. 1e and f) concentration patterns closely matched the titre results from infection assays, demonstrating that SLFN11 diminishes the number of viral particles released from the cells.

To assess a possible reduction of intracellular vRNA, we determined its levels in 293T cells expressing chloramphenicol acetyl transferase (CAT), SLFN5, SLFN11, SLFN11-N or SLFN11-C (Fig. 1g), as well as in 293, 293shRNA$^{\text{Ctl}}$ and 293shRNA$^{\text{SLFN}}$ cells (Fig. 1h). In contrast to the pronounced variations in extracellular vRNA, only insignificant differences in intracellular vRNA were evident among those cells. We also analysed vRNA in the cytoplasmic fraction specifically, as nuclear export of unspliced vRNA is a hallmark of retroviral RNA processing[6–8], but again found no significant alterations attributable to SLFN11 (Supplementary Fig. 4a).

Unlike in BST2-expressing cells[9,10], electron-microscopic analysis failed to document accumulation of virions inside or on the surface of virus-producing cells in the presence of SLFN11 (Supplementary Fig. 4b). As such, SLFN11 greatly diminishes the formation of viral particles inside the cell, despite the fact that vRNA is equally available.

---

[1]Section of Molecular Biology, Division of Biological Sciences, University of California San Diego, La Jolla, California 92093, USA. [2]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, USA. [3]Laboratory of Genetics, The Salk Institute for Biological Studies, La Jolla, California 92037, USA. [4]Moores Cancer Center, University of California San Diego, La Jolla, California 92093, USA.

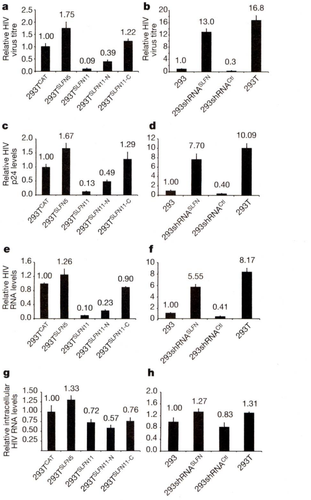


Figure 1 | SLFN11 inhibits retrovirus production without affecting intracellular vRNA levels. a–h, 293T cells were transfected with pNL4-3.Luc.R$^+$E$^-$/pCMV-VSV-G together with SLFN5, SLFN11, SLFN11-N, SLFN11-C or CAT (**a**, **c**, **e**, **g**), or 293, 293shRNA$^{\mathrm{Ctl}}$, 293shRNA$^{\mathrm{SLFN}}$ and 293T cells were transfected with pNL4-3.Luc.R$^+$E$^-$ and pCMV-VSV-G (**b**, **d**, **f**, **h**). **a**, **b**, VSV-G-pseudotyped HIV production was assayed by titrated infection and luciferase assay. **c**, **d**, Viral particle content in supernatants was analysed by p24 ELISA. **e**, **f**, extracellular vRNA concentration was analysed by qPCR of p24. **g**, **h**, intracellular vRNA was determined by qPCR of p24. (average ± s.d.; $n = 3$).

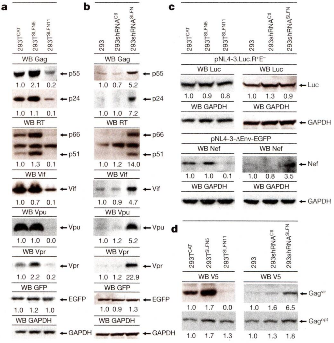


Figure 2 | SLFN11 selectively inhibits viral protein expression on the basis of codon usage. a, 293T cells were transfected with pNL4-3.Luc.R$^+$E$^-$ and pcDNA5-EGFP together with SLFN5, SLFN11 or CAT, and cell lysates immunoblotted for HIV proteins, EGFP and GAPDH. **b**, 293, 293shRNA$^{\mathrm{Ctl}}$ and 293shRNA$^{\mathrm{SLFN}}$ cells were transfected with pNL4-3.Luc.R$^+$E$^-$ and pcDNA5-EGFP, and lysates analysed as in **a**. **c**, Top: 293T cells were co-transfected with pNL4-3.Luc.R$^+$E$^-$ and SLFN5, SLFN11 or CAT (left), or 293, 293shRNA$^{\mathrm{Ctl}}$ and 293shRNA$^{\mathrm{SLFN}}$ cells were transfected with pNL4-3.Luc.R$^+$E$^-$ (right). Lysates were probed for luciferase and GAPDH. Bottom: as above, except pNL4-3-ΔEnv-EGFP was used instead of pNL4-3.Luc.R$^+$E$^-$, and lysates were immunoblotted for Nef and GAPDH; **d**, Cells were transfected with viral codon-usage-based *gag* (Gag$^{\mathrm{vir}}$, top) or synonymous human codon usage-optimized *gag* (Gag$^{\mathrm{opt}}$, bottom). Expression of Gag in cell lysates was determined by anti-V5 immunoblotting.

When the two cell sets were analysed for the expression of viral proteins encoded by pNL4-3.Luc.R$^+$E$^-$, we noted a marked effect of SLFN11 on p55 Gag and p24 capsid proteins[11] (Fig. 2a, b). To create a clearer picture of the effect of SLFN11 on p55, we abolished expression of the viral protease by introducing a stop codon into *pro* in pNL4-3.Luc.R$^+$E$^-$. As anticipated, p24 could no longer be detected; however, modulation of p55 expression by SLFN11 was clearly present (Supplementary Fig. 5a). As with *gag*-derived proteins, SLFN11 had a notable effect on the protein levels of RT, Vif, Vpu and Vpr (Fig. 2a, b). Intriguingly, we did not observe any reduction of enhanced green fluorescent protein (EGFP) derived from a co-transfected vector, or of GAPDH (Fig. 2a, b, bottom), indicating that the limited viral protein production is not due to a global shutdown of protein synthesis.

Even more surprising was that luciferase expression, coded in pNL4-3.Luc.R$^+$E$^-$ in place of *nef*, was mostly impervious to SLFN11 (Fig. 2c, top), contrasting the SLFN11-mediated diminishment of other proteins encoded on this vector. Notably, Nef expression was inhibited by SLFN11 in the context of pNL4-3-ΔEnv-EGFP, which has part of *env* replaced by EGFP, but retains *nef* in its original position. (Fig. 2c, bottom). We thus conclude that SLFN11 selectively suppresses viral protein expression via transcript-intrinsic properties rather than external factors or positional elements. This notion is further corroborated by the fact that SLFN11 had no effect on Rev response element (RRE)-mediated events such as nuclear export of unspliced vRNA (Supplementary Fig. 5b).

Viral genomes have biased nucleotide compositions different from human genes[12–18]. Extremely high frequencies of A nucleotides are found in the RNA genomes of lentiviruses and influenza virus[13,17,19,20]. Wild isolates of HIV-1, particularly *gag* and *pol* sequences, are characterized by low GC content and suboptimal codon usage compared to the host cell preference[14,21–24]. The unusual rare codon bias favours A/U in the third position, which induces ribosome pausing and inefficient translation. As the inhibitory effect of SLFN11 on viral protein expression is intrinsic to the transcripts, we proposed that SLFN11 exploits viral codon preferences to specifically attenuate viral protein synthesis. We therefore generated vectors containing only the open reading frame of HIV-1 *gag* with either viral codon-bias (Gag$^{\mathrm{vir}}$), or synonymous substitutions optimizing for human cell expression (Gag$^{\mathrm{opt}}$). As shown in Fig. 2d, SLFN11 strongly affected expression of Gag$^{\mathrm{vir}}$, but was without consequence for Gag$^{\mathrm{opt}}$ expression. Differences in translation initiation are not likely, as both Gag$^{\mathrm{vir}}$ and Gag$^{\mathrm{opt}}$ contain the same translation start sequences. This finding strongly indicates that SLFN11 is exploiting the distinct viral codon bias to selectively attenuate the expression of viral proteins.

Previous reports indicated changes in cellular tRNA levels after HIV infection[20], prompting us to investigate whether SLFN11 alters the tRNA composition in the absence or presence of HIV. Using tRNA arrays[25,26], we observed little to no changes in tRNA levels as a consequence of SLFN11 expression (Fig. 3a, left column). However, whereas HIV triggered substantial changes in tRNA concentrations in SLFN11-knockdown cells, no changes were observed in the presence of SLFN11 (Fig. 3a, middle and right). Thus, SLFN11 counteracts HIV-induced changes in tRNA composition, presumably initiated to promote viral protein synthesis. To test whether SLFN11 interacts directly with tRNA, we used human tRNA as electrophoretic mobility shift assay (EMSA) probe with fast protein liquid chromatography (FPLC)-purified His-conjugated SLFN11-N. As shown in Fig. 3b, SLFN11-N produced a clear shift of the tRNA in a dose-dependent manner. The shifted band was competed with unlabelled tRNA, but not *in-vitro*-transcribed vRNA (Fig. 3b), and was disrupted by an antibody against SLFN11 (Fig. 3c). To determine whether SLFN11 has binding preference for particular tRNAs, we performed EMSAs with increasing amounts of SLFN11-N to obtain approximately 50% shifting efficiency (Supplementary Fig. 6). When the shifted (S), unshifted (U) and total tRNA (T) bands were isolated and recovered tRNAs hybridized to tRNA arrays, no enrichment of particular tRNAs in the shifted band was noticeable (Supplementary Fig. 6), indicating a lack of discerning binding preference of SLFN11-N.

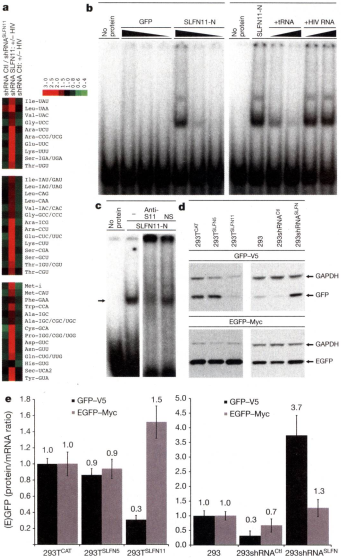


Figure 3 | SLFN11 binds tRNAs and selectively inhibits protein expression based on codon usage. **a**, 293shRNA^{Ctl} and 293shRNA^{SLFN} cells were transfected with pNL4-3.Luc.R$^+$E$^-$ or control vector (indicated as +/− HIV). Relative abundances of mature tRNA species were analysed by microarray as described[25,26]. **b**, Left: increasing amounts of SLFN11-N or GFP were incubated with ^{32}P-labelled tRNA and subjected to EMSA. Right: 2 × or 10 × unlabelled tRNA or *in-vitro*-transcribed vRNA corresponding to the *gag-pol* frame-shifting sequence (120 bases) were added to the binding reaction. **c**, As in **b**, except non-specific (NS) or anti-SLFN11 (Anti-S11) monoclonal antibody was added to the binding reaction. **d**, 293T cells were transfected with V5-tagged GFP, Myc-tagged EGFP and pNL4-3.Luc.R$^+$E$^-$ together with either CAT, SLFN5 or SLFN11 (left). Lysates were probed for V5–GFP, Myc–EGFP and GAPDH. Similarly, V5-tagged GFP, Myc-tagged EGFP and pNL4-3.Luc.R$^+$E$^-$ were co-transfected into 293, 293shRNA^{Ctl} and 293shRNA^{SLFN} cells (right), and V5–GFP, Myc–EGFP and GAPDH expression determined by immunoblotting. **e**, V5–GFP and Myc–EGFP protein levels determined from **d** were quantified and normalized to *V5-GFP* and *Myc-EGFP* mRNA levels, respectively (average ± s.d.; $n = 4$).

The results shown lead to the premise that any protein that uses similar codon usage as HIV would be modulated by HIV and/or SLFN11. To demonstrate unequivocally that codon-bias rather than cryptic regulatory elements in *gag* accounted for the influence of SLFN11, we tested this hypothesis using non-viral proteins. Natural green fluorescent protein (GFP) harbours a similar codon-bias as that of HIV, whereas EGFP has been optimized for mammalian expression by substituting synonymous codons of highly expressed human genes throughout the GFP open reading frame[22]. As anticipated, co-transfection of pNL4-3.Luc.R$^+$E$^-$ with V5-tagged GFP and Myc-tagged EGFP in the absence of SLFN11 resulted in increased GFP, but not EGFP, expression (not shown). Most importantly, SLFN11 inhibited the expression of GFP in a manner identical to viral proteins, whereas EGFP was largely refractory to suppression by SLFN11 (Fig. 3d, e). As (E)GFP protein levels have been normalized to their respective messenger RNA levels (Fig. 3e), we conclude that any differences in GFP expression are the consequence of altered protein synthesis rather than variations in transcription or mRNA stability.

Finally, to demonstrate that the anti-retroviral effects of SLFN11 are not limited to a system using HIV$^{VSV\text{-}G}$ and 293 cells, we used CEM cells, a T-cell line widely used to assess HIV replication kinetics. Human peripheral blood mono-nuclear cells (PBMCs) display similar levels of SLFN11 after IFN-β stimulation as CEM cells (Fig. 4a). Control or SLFN11-directed shRNAs were used to generate stable CEM variants (Fig. 4b), and the resulting CEMshRNA^{Ctl} and CEMshRNA^{SLFN} cells and parental CEM cells were infected with wild-type X4-HIV-1$_{LAI}$, and viral replication was assessed via p24 enzyme-linked immunosorbent assays (ELISA) of the supernatants.

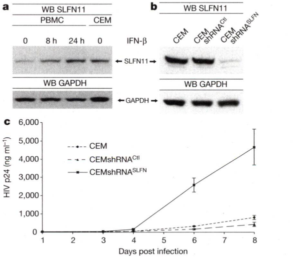


Figure 4 | SLFN11 inhibits replication of wild-type HIV-1$_{LAI}$ in CEM cells. **a**, Human PBMCs were stimulated with 6,000 U ml^{-1} IFN-β. SLFN11 expression in the derived lysates and in CEM cell lysates was analysed by immunoblotting. **b**, SLFN11 expression in CEM, CEMshRNA^{Ctl} and CEMshRNA^{SLFN} cells as analysed by immunoblotting. **c**, CEM, CEMshRNA^{Ctl} and CEMshRNA^{SLFN} cells were infected with HIV-1$_{LAI}$ at a multiplicity of infection (m.o.i.) of 0.01, and viral replication was assayed by p24 ELISA of culture supernatants (average ± s.d.; $n = 4$).

As anticipated, CEM and $\text{CEMshRNA}^{\text{Ctl}}$ cells produced comparable viral titres at all time points. In contrast, $\text{CEMshRNA}^{\text{SLFN}}$ cells facilitated significantly enhanced HIV-1 replication, yielding an exponentially increasing difference in viral titres (Fig. 4c). A second $\text{CEMshRNA}^{\text{SLFN}}$ cell line established by means of a distinct *SLFN11* shRNA further corroborated these results (Supplementary Fig. 7).

In summary, systematic analysis of the HIV replication cycle revealed that SLFN11 does not inhibit reverse transcription, integration or production and nuclear export of viral RNA, nor did we observe a block in budding or release of viral particles. However, we discovered a selective inhibition in the synthesis of virally encoded proteins. Intriguingly, a specific inhibition of viral protein synthesis in HIV-infected cells in response to interferon was previously observed, but the factors responsible for the effect were not identified[27]. We postulate that SLFN11 acts at the point of virus protein synthesis by exploiting the unique viral codon bias towards A/T nucleotides. This model supports the notion that the antiviral activity of SLFN11 extends to other viruses with rare codon bias such as influenza, but apparently not to AAV or HSV. The exact mechanism by which HIV alters tRNA function in its favour and how SLFN11 counteracts this process will require considerable further analysis. Evidently, SLFN11 interacts with all tRNAs *in vitro*. Direct binding of SLFN11 to tRNA offers the possibility that SLFN11 either sequesters tRNAs, prevents their maturation via post-transcriptional processing or accelerates their deacylation. In either case, if already rare tRNAs are further reduced, tRNA availability might manifest as the rate-limiting step in the synthesis of proteins involving those tRNAs. In contrast, a lesser or no impact would be expected on proteins synthesized via plentiful tRNAs, as even if a fraction of those tRNAs is 'eliminated', translation initiation will likely remain the rate-limiting event. In conclusion, our results establish SLFN11 as a potent, interferon-inducible restriction factor against retroviruses such as HIV, mediating its antiviral effects on the basis of codon usage discrimination.

METHODS SUMMARY

Plasmids and antibodies. SLFN5 and SLFN11 were cloned into pcDNA6/V5-His. PCR fragments of SLFN11 were cloned into pcDNA6/V5-His to generate SLFN11-N (amino acids 1–579) and SLFN11-C (amino acids 523–901). The pNL4-3.Luc.R$^+$E$^-$ HIV-1 vector has been described previously[28]. SLFN11 antibodies were from Sigma and Abmart. pNL4-3-ΔEnv-EGFP, pNL-GFP-RRE(SA) and antibodies against HIV-1 proteins were obtained through the NIH AIDS Research and Reference Reagent Program. Generation of Gag^{vir}, Gag^{opt} and (E)GFP vectors is in Methods. pLKO.1 shRNA lentivirus constructs were from Open Biosystems.

Virus production and titre assays. HIV production was determined by transfection of pNL4-3.Luc.R$^+$E$^-$ and pCMV-VSV-G into cells of interest. Supernatants were used to spin-infect HEK293T cells, and luciferase was measured by Bright-Glo Luciferase Assay (Promega). p24 ELISAs were performed at the Center for AIDS Research, UCSD using Alliance p24 ELISA kits (Perkin Elmer).

HIV replication kinetics. Wild type X4 strain of HIV-1_{LAI} was used to inoculate CEM, $\text{CEMshRNA}^{\text{Ctl}}$ and $\text{CEMshRNA}^{\text{SLFN}}$ cells at a multiplicity of infection of 0.01, and viral titres were measured by p24 ELISA.

Data analysis and presentation. Unless indicated otherwise, results presented in graphs are average ± s.d. of at least three independent transfections or infections. For immunoblots, a representative of at least three independent transfections or infections is shown.

Full Methods and any associated references are available in the online version of the paper.

Received 22 November 2011; accepted 23 July 2012.

Published online 23 September 2012.

1. Sohn, W. J. *et al.* Novel transcriptional regulation of the schlafen-2 gene in macrophages in response to TLR-triggered stimulation. *Mol. Immunol.* **44,** 3273–3282 (2007).
2. Yoneyama, M. *et al.* Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity. *J. Immunol.* **175,** 2851–2858 (2005).
3. Schwarz, D. A., Katayama, C. D. & Hedrick, S. M. *Schlafen*, a new family of growth regulatory genes that affect thymocyte development. *Immunity* **9,** 657–668 (1998).
4. Geserick, P., Kaiser, F., Klemm, U., Kaufmann, S. H. & Zerrahn, J. Modulation of T cell development and activation by novel members of the Schlafen (slfn) gene family harbouring an RNA helicase-like motif. *Int. Immunol.* **16,** 1535–1548 (2004).
5. Berger, M. *et al.* An *Slfn2* mutation causes lymphoid and myeloid immunodeficiency due to loss of immune cell quiescence. *Nature Immunol.* **11,** 335–343 (2010).
6. Fritz, C. C. & Green, M. R. HIV Rev uses a conserved cellular protein export pathway for the nucleocytoplasmic transport of viral RNAs. *Curr. Biol.* **6,** 848–854 (1996).
7. Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W. & Luehrmann, R. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82,** 475–483 (1995).
8. Malim, M. H., Hauber, J., Le, S. Y., Maizel, J. V. & Cullen, B. R. The HIV-1 rev *trans*-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **338,** 254–257 (1989).
9. Van Damme, N. *et al.* The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. *Cell Host Microbe* **3,** 245–252 (2008).
10. Neil, S. J., Zang, T. & Bieniasz, P. D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451,** 425–430 (2008).
11. Ganser-Pornillos, B. K., Yeager, M. & Sundquist, W. I. The structural biology of HIV assembly. *Curr. Opin. Struct. Biol.* **18,** 203–217 (2008).
12. Coleman, J. R. *et al.* Virus attenuation by genome-scale changes in codon pair bias. *Science* **320,** 1784–1787 (2008).
13. Zhou, T., Gu, W., Ma, J., Sun, X. & Lu, Z. Analysis of synonymous codon usage in H5N1 virus and other influenza A viruses. *Biosystems* **81,** 77–86 (2005).
14. Meintjes, P. L. & Rodrigo, A. G. Evolution of relative synonymous codon usage in human immunodeficiency virus type-1. *J. Bioinform. Comput. Biol.* **3,** 157–168 (2005).
15. Frelin, L. *et al.* Codon optimization and mRNA amplification effectively enhances the immunogenicity of the hepatitis C virus nonstructural 3/4A gene. *Gene Ther.* **11,** 522–533 (2004).
16. Forsberg, R. & Christiansen, F. B. A codon-based model of host-specific selection in parasites, with an application to the influenza A virus. *Mol. Biol. Evol.* **20,** 1252–1259 (2003).
17. Plotkin, J. B. & Dushoff, J. Codon bias and frequency-dependent selection on the hemagglutinin epitopes of influenza A virus. *Proc. Natl Acad. Sci. USA* **100,** 7152–7157 (2003).
18. Grantham, P. & Perrin, P. AIDS virus and HTLV-I differ in codon choices. *Nature* **319,** 727–728 (1986).
19. Wong, E. H. M., Smith, D. K., Rabadan, R., Peiris, M. & Poon, L. L. M. Codon usage bias and the evolution of influenza A viruses. Codon usage biases of influenza virus. *BMC Evol. Biol.* **10,** 253 (2010).
20. van Weringh, A. *et al.* HIV-1 modulates the tRNA pool to improve translation efficiency. *Mol. Biol. Evol.* **28,** 1827–1834 (2011).
21. Kofman, A. *et al.* HIV-1 gag expression is quantitatively dependent on the ratio of native and optimized codons. *Tsitologiia* **45,** 86–93 (2003).
22. Haas, J., Park, E. C. & Seed, B. Codon usage limitation in the expression of HIV-1 envelope glycoprotein. *Curr. Biol.* **6,** 315–324 (1996).
23. Berkhout, B. & van Hemert, F. J. The unusual nucleotide content of the HIV RNA genome results in a biased amino acid composition of HIV proteins. *Nucleic Acids Res.* **22,** 1705–1711 (1994).
24. Kypr, J. & Mrazek, J. Unusual codon usage of HIV. *Nature* **327,** 20 (1987).
25. Pavon-Eternod, M. *et al.* tRNA over-expression in breast cancer and functional consequences. *Nucleic Acids Res.* **37,** 7268–7280 (2009).
26. Dittmar, K. A., Mobley, E. M., Radek, A. J. & Pan, T. Exploring the regulation of tRNA distribution on the genomic scale. *J. Mol. Biol.* **337,** 31–47 (2004).
27. Coccia, E. M., Krust, B. & Hovanessian, A. G. Specific inhibition of viral protein synthesis in HIV-infected cells in response to interferon treatment. *J. Biol. Chem.* **269,** 23087–23094 (1994).
28. Connor, R. I., Chen, B. K., Choe, S. & Landau, N. R. Vpr is required for efficient replication of human immunodeficiency virus type-1 in mononuclear phagocytes. *Virology* **206,** 935–944 (1995).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Smith for help with the HIV replication studies; M. Wood for performing the electron microscope analysis; D. Xu and D.-Y. Song for technical assistance; and J. Young, J. Guatelli, D. Smith, M. Kaul and S. Chanda for discussion. This work was supported in part by NIH AI81019 and AI074967 to M.D.W., NIH P01AI090935, R01GM101982 and R21AI088490 to M.D., and by resources from the UCSD Center for AIDS Research, NIH P30AI36214, and the HINT Program, NIH P01AI090935. The authors declare no competing financial interests.

Author Contributions M.L., M.D.W., T.P. and M.D. planned the experiments; M.L., E.K., X.G., M.P.-E., K.L., H.S., T.E.J. and S.L. conducted the experiments; M.L., T.P., M.D.W. and M.D. analysed the data; and M.L. and M.D. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.D. (midavid@ucsd.edu).

LETTER

doi:10.1038/nature11443

Vaccine-induced CD8$^+$ T cells control AIDS virus replication

Philip A. Mudd[1,2], Mauricio A. Martins[3], Adam J. Ericsen[1], Damien C. Tully[4], Karen A. Power[4], Alex T. Bean[1], Shari M. Piaskowski[1], Lijie Duan[5], Aaron Seese[4], Adrianne D. Gladden[4], Kim L. Weisgrau[1], Jessica R. Furlott[1], Young-il Kim[6], Marlon G. Veloso de Santana[7], Eva Rakasz[8], Saverio Capuano III[8], Nancy A. Wilson[1,8], Myrna C. Bonaldo[7], Ricardo Galler[9], David B. Allison[10], Michael Piatak Jr[11], Ashley T. Haase[5], Jeffrey D. Lifson[11], Todd M. Allen[4] & David I. Watkins[3]

Developing a vaccine for human immunodeficiency virus (HIV) may be aided by a complete understanding of those rare cases in which some HIV-infected individuals control replication of the virus[1–3]. Most of these elite controllers express the histocompatibility alleles *HLA-B*57* or *HLA-B*27* (ref. 3). These alleles remain by far the most robust associations with low concentrations of plasma virus[4,5], yet the mechanism of control in these individuals is not entirely clear. Here we vaccinate Indian rhesus macaques that express *Mamu-B*08*, an animal model for HLA-B*27-mediated elite control[6], with three Mamu-B*08-restricted CD8$^+$ T-cell epitopes, and demonstrate that these vaccinated animals control replication of the highly pathogenic clonal simian immunodeficiency virus (SIV) mac239 virus. High frequencies of CD8$^+$ T cells against these Vif and Nef epitopes in the blood, lymph nodes and colon were associated with viral control. Moreover, the frequency of the CD8$^+$ T-cell response against the Nef RL10 epitope (Nef amino acids 137–146) correlated significantly with reduced acute phase viraemia. Finally, two of the eight vaccinees lost control of viral replication in the chronic phase, concomitant with escape in all three targeted epitopes, further implicating these three CD8$^+$ T-cell responses in the control of viral replication. Our findings indicate that narrowly targeted vaccine-induced virus-specific CD8$^+$ T-cell responses can control replication of the AIDS virus.

A vaccine is desperately needed to curb the global HIV pandemic. Estimates project that for every HIV-infected individual initiating antiretroviral treatment, more than two individuals are newly infected[7]. Vaccines have historically been chosen based on their ability to induce responses that mimic successful immune responses to human pathogens, yet correlates of successful immune responses against HIV remain an enigma. Elite control of chronic phase viral replication is the best example of an effective immune response against the virus[1]. Understanding why elite controllers control viral replication may enable the design of an effective vaccine for HIV.

Recently, we have discovered similarities between an animal model of elite control and the same phenomenon in humans[8]. The Indian rhesus macaque major histocompatibility complex (MHC) class I molecule Mamu-B*08 and the human leukocyte antigen (HLA) class I molecule HLA-B*27 bind many of the same peptides[6]. Despite being divergent at 28 amino acid positions, these molecules share similar peptide-binding motifs, including an identical position 2 arginine primary anchor[6]. Both MHC class I molecules also exhibit considerable overlap in preferred binding residues at the other dominant position 1 and position 9 residues[6]. Remarkably, 50% of *Mamu-B*08*$^+$ animals show some measure of control of the highly pathogenic SIVmac239 virus[6,9]. It is important to note, however, that several key differences exist between HLA-B*27-mediated elite control and the *Mamu-B*08*$^+$ animal model. Humans are infected with a variety of different viruses and it is widely thought that CD8$^+$ T cells directed against conserved epitopes in Gag have an important role in viral control[1]. By contrast, the animal model of elite control has been developed using macaques infected only with the SIVmac239 clone. Furthermore, *Mamu-B*08*$^+$ macaques do not usually recognize Gag-derived epitopes[10]. Finally, Indian rhesus macaques infected with SIVmac239 have average chronic phase viral loads of more than 10^6 viral RNA (vRNA) copies per millilitre, whereas humans normally have mean plasma viral concentrations of 30,000 vRNA per millilitre in the chronic phase. Owing to the high replicative capacity of SIVmac239, most drug- or vaccine-naive SIVmac239-infected monkeys die from AIDS-defining illnesses by 2 years after infection[11].

Three Mamu-B*08-restricted CD8$^+$ T-cell responses make up more than half of the T-cell response against the virus in *Mamu-B*08*$^+$ elite controllers[10,12]; these CD8$^+$ T cells recognize Vif RL8 (Vif amino acids 172–179), Vif RL9 (Vif amino acids 123–131) and Nef RL10 (Nef amino acids 137–146). CD8$^+$ T cells directed against these three epitopes select for escape mutations preferentially in *Mamu-B*08*$^+$ macaques that do not become elite controllers[13], suggesting that these three immunodominant T-cell responses are important for the development of elite control.

To test the hypothesis that antigen-specific CD8$^+$ T cells are responsible for the development of elite control, we vaccinated eight *Mamu-B*08*$^+$ macaques (Fig. 1a, group 1) with two small regions of the SIV proteome that include the three immunodominant T-cell epitopes bound by Mamu-B*08: Vif RL8, Vif RL9 and Nef RL10. We used a recombinant yellow fever 17D (rYF17D) prime and boosted with recombinant adenovirus serotype 5 (rAd5). As controls, we vaccinated another eight *Mamu-B*08*$^+$ macaques (Fig. 1a, group 2) with two small regions of the SIV proteome that do not encode any known Mamu-B*08 epitopes[10]. The group 1 vaccinees mounted high frequency CD8$^+$ T-cell responses against the three Mamu-B*08-bound epitopes after vaccination, whereas the group 2 animals did not make CD8$^+$ T-cell responses against these three epitopes (Fig. 1b). Group 1 and 2 animals exhibited similar SIV-specific total T-cell and CD4$^+$ T-cell response frequencies against the vaccine inserts after vaccination, as demonstrated by interferon (IFN)-γ enzyme-linked immunospot (ELISPOT) assay (Supplementary Fig. 1a, b).

Fifteen weeks after the final vaccine boost, we challenged all 16 animals intrarectally with a high dose of SIVmac239 (10,000 half-maximal tissue-culture infectious dose ($TCID_{50}$)). Four out of eight group 1 animals and six out of eight group 2 animals were infected after this first challenge. The remaining uninfected macaques were

---

[1]Department of Pathology and Laboratory Medicine, University of Wisconsin-Madison, Madison, Wisconsin 53711, USA. [2]Medical Scientist Training Program, University of Wisconsin-Madison, Madison, Wisconsin 53705, USA. [3]Department of Pathology, University of Miami Miller School of Medicine, Miami, Florida 33136, USA. [4]Ragon Institute of MGH, MIT and Harvard, Boston, Massachusetts 02129, USA. [5]Department of Microbiology, University of Minnesota, Minneapolis, Minnesota 55455, USA. [6]Department of Medicine, Division of Preventive Medicine, University of Alabama at Birmingham, Birmingham, Alabama 35294, USA. [7]Laboratório de Biologia Molecular de Flavivírus, Instituto Oswaldo Cruz—FIOCRUZ, Rio de Janeiro, Brazil. [8]Wisconsin National Primate Research Center, University of Wisconsin-Madison, Madison, Wisconsin 53711, USA. [9]Instituto de Tecnologia em Imunobiológicos, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil. [10]Department of Biostatistics, Section on Statistical Genetics, University of Alabama at Birmingham, Birmingham, Alabama 35294, USA. [11]AIDS and Cancer Virus Program, SAIC Frederick, Inc., National Cancer Institute, Frederick, Maryland 21702, USA.

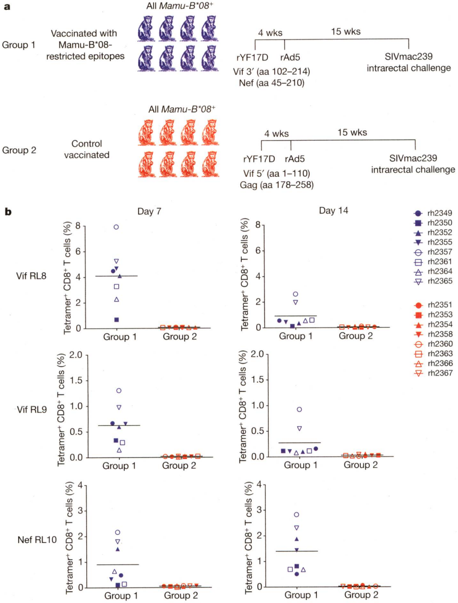


Figure 1 | Experimental design. **a**, *Mamu-B*08*$^+$ macaques were vaccinated with an identical rYF17D/rAd5 regimen. Group 1 animals received two SIVmac239 constructs: Vif 3′ (amino acids (aa) 102–214; includes the Mamu-B*08-restricted Vif$_{123-131}$RL9 and Vif$_{172-179}$RL8 epitopes) and Nef (amino acids 45–210; includes the Mamu-B*08-restricted Nef$_{137-146}$RL10 epitope). Group 2 received two SIVmac239 constructs: Gag (amino acids 178–258) and Vif 5′ (amino acids 1–110), which do not contain any known Mamu-B*08-restricted CD8$^+$ T-cell epitopes[10]. Animals were challenged with SIVmac239 intrarectally 15 weeks after the rAd5 boost. **b**, Frequencies of tetramer$^+$ CD8$^+$ T cells in PBMCs at days 7 and 14 after the rAd5 boost. Lines represent mean frequency.

rechallenged with a second high dose of SIVmac239 (10,000 TCID$_{50}$) three weeks after the initial challenge. All of the group 2 control-vaccinated animals were infected after this second challenge. A single group 1 vaccinee, rh2355, resisted four high-dose SIVmac239 challenges and was only infected after the fifth challenge.

Surprisingly, all eight infected group 1 animals, vaccinated with the three immunodominant Mamu-B*08-restricted T-cell epitopes, controlled viral replication during acute infection when compared with the group 2 animals (Fig. 2a–d). Remarkably, six of the seven group 1 vaccinated *Mamu-B*08*$^+$ animals that were infected after one or two challenges became elite controllers with post-acute phase viral loads of less than 1,000 vRNA copies per ml (Fig. 2a). The seventh animal (rh2349) controlled viral replication to less than 10,000 vRNA copies per ml by 8 weeks after infection. The group 1 animal (rh2355) that required five high-dose challenges for infection controlled viral replication to less than 10,000 vRNA copies per ml by 10 weeks after infection. By contrast, viral replication in the group 2 animals was indistinguishable from viral replication in four concurrently challenged unvaccinated *Mamu-B*08*$^+$ controls (Fig. 2c). We also measured viral replication during acute infection in lymph node biopsies using *in situ* hybridization[14,15]. Group 1 animals demonstrated significantly less viral replication in lymph node tissue at time points near peak plasma viral load than group 2 animals (Fig. 2e).

To evaluate potential correlates of viral protection in the group 1 vaccinees, we characterized the kinetics of T-cell responses after SIV infection using peptides and Mamu-B*08 tetramers. We first measured SIV-specific T-cell responses after infection in both group 1 and 2 vaccinated animals with a panel of synthetic peptides spanning

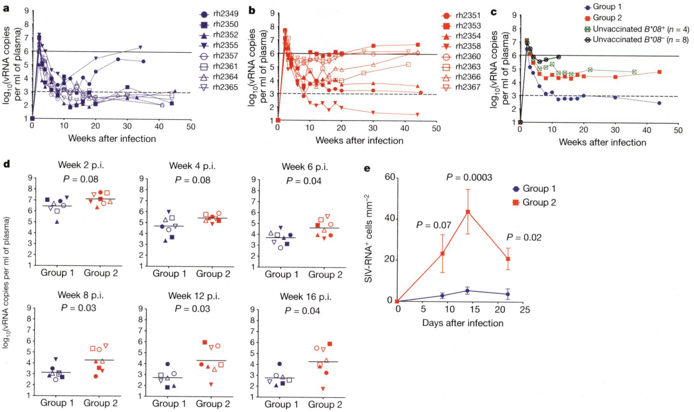


Figure 2 | *In vivo* viral replication. **a**, Plasma viral loads (VL) for group 1 animals. The dashed line (10^3 vRNA copies per ml) represents the threshold that defines elite control in Indian rhesus macaques[9]. **b**, Plasma viral loads for group 2 control animals. **c**, Geometric mean plasma viral loads for groups 1, 2 and unvaccinated *Mamu-B*08*$^+$ ($n = 4$) or *Mamu-B*08*$^-$ ($n = 8$) control macaques. **d**, Plasma viral load comparisons between groups 1 and 2. Lines represent geometric means. p.i., post-infection. **e**, Viral replication measured using *in situ* hybridization[14,15] in lymph node biopsy specimens. Data are mean ± s.e.m.

both the vaccine inserts and the entire SIV proteome using IFN-γ ELISPOT (Supplementary Fig. 1a, b). Interestingly, peripheral blood mononuclear cell (PBMC) responses against the entire proteome were equivalent in the two groups after infection. At days 14 and 17 after infection, however, higher frequency insert-specific (total T cell and CD4$^+$ T-cell) and proteome-specific (CD4$^+$ T-cell) responses were observed in the group 1 animals, perhaps reflecting preservation of the CD4$^+$ T-cell responses owing to reduced peak viral replication in this group. These preserved antigen-specific CD4$^+$ T-cell responses may have facilitated the control of chronic phase viral replication in the group 1 animals. We also found that group 1 vaccinated macaques exhibited earlier and higher magnitude Vif RL9- and Nef RL10-specific responses in the peripheral blood than the group 2 macaques (Fig. 3a). This pattern was also detected in lymph node biopsy specimens (Fig. 3b) and colon biopsy tissue (Fig. 3c). Notably, T-cell responses directed against the Vif RL8 epitope were similar between the group 1 and 2 macaques (Fig. 3a–c). Taken together, these results suggest that early, high frequency Vif RL9- and Nef RL10-specific responses in the blood and in key immunological tissues were directly responsible for viral control in the group 1 *Mamu-B*08*$^+$ macaques.

Remarkably, we found a significant correlation between high frequency Nef RL10-specific responses and decreased day 14 peak viral loads (Fig. 3d) in our cohort of infected macaques. Furthermore, these vaccine-induced Nef RL10-specific CD8$^+$ T cells expressed both granzyme B and perforin at day 17 after infection, at a time when Nef RL10-specific CD8$^+$ T cells showed little or no expression of these important markers of cytotoxicity in the group 2 animals (Fig. 3e, f and Supplementary Fig. 2). Furthermore, the frequency of Nef RL10- (but not Vif RL8- or Vif RL9-) specific CD8$^+$ T-cell responses has recently been shown to correlate with the control of chronic phase viral replication in SIVmac239-infected, unvaccinated *Mamu-B*08*$^+$ macaques (D. Douek, personal communication). These results suggest that early, high frequency cytotoxic Nef RL10-specific responses in the blood and in key immunological tissues were directly responsible for viral control in group 1 *Mamu-B*08*$^+$ macaques. This is the first described correlate of viral control in the Mamu-B*08 model of elite control.

After controlling viral replication during the acute phase, increased plasma viraemia was observed in two of the group 1 vaccinees during the chronic phase (rh2349 and rh2355). Viral sequencing showed that despite the presence of wild-type virus in one of these two animals at 6 weeks after infection, all three targeted epitopes had escaped by the time viral replication started to increase in the chronic phase (Table 1 and Supplementary Fig. 3a, b). By contrast, most of the three epitopes were intact in virus from two of the elite controllers in group 1 at 48 weeks after infection (Table 1 and Supplementary Fig. 3b). This finding further implicates CD8$^+$ T cells against these three epitopes in control of viral replication, and suggests that the virus cannot replicate in the face of these CD8$^+$ T-cell responses. The canonical escape mutants that arose in these animals did not seem to affect viral fitness because the virus replicated to high levels in these two 'breakthrough' animals. Furthermore, we have previously engineered mutant viruses bearing several of the escape mutants observed in these animals and these mutations caused abrogation of CD8$^+$ T-cell recognition and did not seem to cause significant fitness loss *in vitro* and *in vivo*[10].

With this report, we demonstrate that vaccine-induced Vif- and Nef-specific CD8$^+$ T cells can control replication of a highly pathogenic

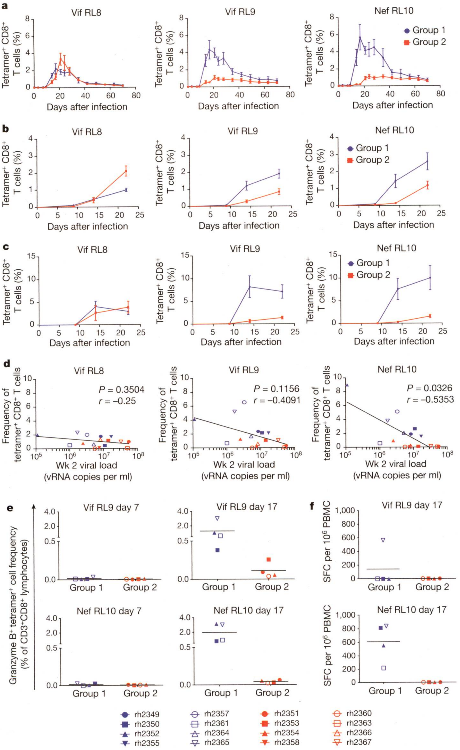


Figure 3 | CD8$^+$ T-cell responses in vaccinated animals after SIVmac239 infection. **a–c**, Peptide/Mamu-B*08 tetramer staining in PBMCs (**a**), lymph node biopsies (**b**) and colon biopsies (**c**). Error bars represent s.e.m. **d**, Correlations between the frequency of tetramer$^+$ cells and the viral load at week 2 after infection. **e**, Granzyme B production by SIV-specific CD8$^+$ T cells in group 1 and 2 vaccinees at days 7 and 17 after SIV infection. **f**, Perforin production by SIV-specific CD8$^+$ T cells in group 1 and 2 vaccinees at day 17 after infection. Assay results are shown as spot forming cells (SFC) per 10^6 PBMCs.

AIDS virus in an animal model of MHC class I-associated elite control. Further analysis of the Vif RL9- and Nef RL10-specific CD8$^+$ T-cell responses in *Mamu-B*08*$^+$ macaques could shed light on the general principles of effective AIDS virus-specific CD8$^+$ T-cell responses. It is possible that vaccines that generate high frequency efficacious effector CD8$^+$ T cells against only a few epitopes in the mucosae might intercept the virus during this crucial acute phase and result in long-term virological control. These types of effective immune response might be inducible in the setting of an appropriate vaccination regimen[16–19]. Therefore, understanding why these particular T-cell responses control viral replication when most other T-cell responses do not may enable the design of an effective approach to HIV vaccination.

Table 1 | Viral sequence analysis

	Vif RL8	Vif RL9	Nef RL10
Acute phase, group 1 animals			
rh2349 (wk 6)	**ND**	**ND**	**ND**
rh2355 (wk 6)	**100%**	**100%**	**100%**
rh2350 (wk 6)	99%	27%	ND
rh2352 (wk 6)	97%	47%	ND
rh2357 (wk 6)	87%	75%	97%
rh2361 (wk 6)	100%	100%	100%
rh2364 (wk 6)	59%	37%	92%
rh2365 (wk 6)	ND	ND	ND
Acute phase, group 2 animals			
rh2351 (wk 6)	22%	33%	97%
rh2353 (wk 6)	38%	40%	95%
rh2354 (wk 6)	92%	92%	100%
rh2358 (wk 6)	100%	100%	99%
rh2360 (wk 6)	50%	64%	99%
rh2363 (wk 6)	94%	99%	95%
rh2366 (wk 6)	100%	69%	99%
rh2367 (wk 6)	97%	78%	100%
Chronic phase, group 1 animals			
rh2349 (wk 35)	**4%**	**4%**	**7%**
rh2355 (wk 37)	**0%**	**0%**	**0%**
rh2350 (wk 48)	ND	ND	91%
rh2352 (wk 48)	97%	17%	69%
rh2357 (wk 45)	ND	ND	99%
rh2361 (wk 48)	99%	68%	99%
rh2364 (wk 45)	ND	ND	ND
rh2365 (wk 48)	ND	ND	ND

454 sequencing of the three Mamu-B*08-restricted Vif RL8, Vif RL9 and Nef RL10 CD8$^+$ T-cell epitopes. Vif and Nef were amplified independently and PCR amplicons were sequenced by 454. Sequence analysis was carried out using RC454 and V-phaser as previously described[21,22]. Percentages indicate the frequencies of wild-type epitope sequence reads for both the acute and chronic phases of infection. Breakthrough animals in group 1 are in bold. ND, not determined.

METHODS SUMMARY

Animals. Animals were cared for according to Animal Welfare Assurance no. A3368-01 (protocol no. G00639).

Vaccination. Animals were vaccinated subcutaneously with two separate rYF17D constructs (2.0×10^5 plaque-forming units of each) and were then boosted intramuscularly with two separate rAd5 vectors (10^{11} particles of each) containing the same SIVmac239 inserts.

SIV infection, *in situ* hybridization and viral load measurement. Animals were challenged intrarectally with 10,000 $TCID_{50}$ of SIVmac239. Viral loads were measured from EDTA anti-coagulated plasma using a previously described protocol[20]. *In situ* hybridization of lymph node biopsy tissues was performed as previously described[14,15].

Tetramer staining and flow cytometry. For tetramer staining, we followed previously published staining protocols[10].

Amplicon-based 454 sequencing of Vif and Nef epitopes in plasma virus. We used an amplicon-based 454 sequencing approach to analyse the Vif and Nef epitopes as previously described[13]. A single-step reverse transcription PCR (RT–PCR) was carried out for each of the unique amplicon/animal/multiplex-identifier-tag sequence combinations using SuperScript III one-step RT–PCR system. Sequencing and run processing were performed on a GS Junior 454 sequencing instrument. We used the ReadClean 454 (RC454) and V-Phaser algorithms as previously described[21,22] to call variants from the 454 data sets.

Statistics. We compared the geometric mean of viral loads and the mean of SIV-RNA$^+$ cells per mm^2 for each time point and then performed Mann–Whitney tests. We also used the Mann–Whitney test to compare the magnitude of SIV-specific T-cell responses in groups 1 and 2. We determined correlation coefficients (r) and P values using the Spearman rank correlation test. All significance tests were two-tailed.

Full Methods and any associated references are available in the online version of the paper.

Received 3 April; accepted 26 July 2012.

Published online 30 September 2012.

1. Deeks, S. G. & Walker, B. D. Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. *Immunity* **27,** 406–416 (2007).
2. Goulder, P. J. & Watkins, D. I. HIV and SIV CTL escape: implications for vaccine design. *Nature Rev. Immunol.* **4,** 630–640 (2004).
3. Migueles, S. A. & Connors, M. Long-term nonprogressive disease among untreated HIV-infected individuals: clinical implications of understanding immune control of HIV. *J. Am. Med. Assoc.* **304,** 194–201 (2010).
4. Fellay, J. *et al.* A whole-genome association study of major determinants for host control of HIV-1. *Science* **317,** 944–947 (2007).
5. The International HIV Controllers Study. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science* **330,** 1551–1557 (2010).
6. Loffredo, J. T. *et al.* Two MHC class I molecules associated with elite control of immunodeficiency virus replication, Mamu-B*08 and HLA-B*2705, bind peptides with sequence similarity. *J. Immunol.* **182,** 7763–7775 (2009).
7. Johnston, M. I. & Fauci, A. S. An HIV vaccine–challenges and prospects. *N. Engl. J. Med.* **359,** 888–890 (2008).
8. Mudd, P. A. & Watkins, D. I. Understanding animal models of elite control: windows on effective immune responses against immunodeficiency viruses. *Curr. Opin. HIV AIDS* **6,** 197–201 (2011).
9. Loffredo, J. T. *et al.* *Mamu-B*08*-positive macaques control simian immunodeficiency virus replication. *J. Virol.* **81,** 8827–8832 (2007).
10. Valentine, L. E. *et al.* Infection with "escaped" virus variants impairs control of simian immunodeficiency virus SIVmac239 replication in *Mamu-B*08*-positive macaques. *J. Virol.* **83,** 11514–11527 (2009).
11. Dang, Q. & Hirsch, V. M. Rapid disease progression to AIDS due to Simian immunodeficiency virus infection of macaques: host and viral factors. *Adv. Pharmacol.* **56,** 369–398 (2008).
12. Loffredo, J. T. *et al.* Patterns of CD8$^+$ immunodominance may influence the ability of *Mamu-B*08*-positive macaques to naturally control simian immunodeficiency virus SIVmac239 replication. *J. Virol.* **82,** 1723–1738 (2008).
13. Mudd, P. A. *et al.* Escape from CD8$^+$ T cell responses in Mamu-B*00801$^+$ macaques differentiates progressors from elite controllers. *J. Immunol.* **188,** 3364–3370 (2012).
14. Li, Q. *et al.* Peak SIV replication in resting memory CD4$^+$ T cells depletes gut lamina propria CD4$^+$ T cells. *Nature* **434,** 1148–1152 (2005).
15. Li, Q. *et al.* Glycerol monolaurate prevents mucosal SIV transmission. *Nature* **458,** 1034–1038 (2009).
16. Hansen, S. G. *et al.* Effector memory T cell responses are associated with protection of rhesus monkeys from mucosal simian immunodeficiency virus challenge. *Nature Med.* **15,** 293–299 (2009).
17. Hansen, S. G. *et al.* Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine. *Nature* **473,** 523–527 (2011).
18. Kawada, M. *et al.* Gag-specific cytotoxic T-lymphocyte-based control of primary simian immunodeficiency virus replication in a vaccine trial. *J. Virol.* **82,** 10199–10206 (2008).
19. Tsukamoto, T. *et al.* Impact of cytotoxic-T-lymphocyte memory induction without virus-specific CD4$^+$ T-cell help on control of a simian immunodeficiency virus challenge in rhesus macaques. *J. Virol.* **83,** 9339–9346 (2009).
20. Cline, A. N., Bess, J. W., Piatak, M. J. & Lifson, J. D. Highly sensitive SIV plasma viral load assay: practical considerations, realistic performance expectations, and application to reverse engineering of vaccines for AIDS. *J. Med. Primatol.* **34,** 303–312 (2005).
21. Henn, M. R. *et al.* Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. *PLoS Pathog.* **8,** e1002529 (2012).
22. Macalalad, A. R. *et al.* Highly sensitive and specific detection of rare variants in mixed viral populations from massively parallel sequence data. *PLOS Comput. Biol.* **8,** e1002417 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank M. Stevenson, M. Reynolds, N. Maness and J. Sacha for discussions and suggestions. We also thank D. Evans, L. Heyer and Z. R. Bergman for facilitating the experiments. This work was funded in part by National Institutes of Health (NIH) grants R37 AI052056, RO1 AI076114, RR015371, contract number HHSN261200800001E and FAPERJ, INCTV, CNPq, MCT and FIOCRUZ.

Author Contributions D.I.W., T.M.A. and P.A.M. conceived of and designed the experiments. P.A.M., A.J.E., M.A.M., K.A.P., A.S., A.D.G., A.T.B., S.M.P., L.D., K.L.W., J.R.F., S.C., M.P., A.T.H. and J.D.L. performed the experiments. P.A.M., M.A.M., N.A.W., E.R., M.P., A.T.H., J.D.L., D.C.T., T.M.A. and D.I.W. compiled and analysed the data. Y.K. and D.B.A. oversaw statistical analysis of the data and aided in interpretation. M.G.V., M.C.B. and R.G. designed and manufactured the rYF17D vectors used in these studies. P.A.M., A.T.H., J.D.L., T.M.A. and D.I.W. wrote the manuscript.

Author Information The sequence data for this study are available at the NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/Traces/sra) under accession number SRA055739. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.I.W. (dwatkins@med.miami.edu).

LETTER

doi:10.1038/nature11517

The molecular basis of phosphate discrimination in arsenate-rich environments

Mikael Elias[1]*, Alon Wellner[1]*, Korina Goldin-Azulay[1], Eric Chabriere[2], Julia A. Vorholt[3], Tobias J. Erb[3] & Dan S. Tawfik[1]

Arsenate and phosphate are abundant on Earth and have striking similarities: nearly identical pK_a values[1,2], similarly charged oxygen atoms, and thermochemical radii that differ by only 4% (ref. 3). Phosphate is indispensable and arsenate is toxic, but this extensive similarity raises the question whether arsenate may substitute for phosphate in certain niches[4,5]. However, whether it is used or excluded, discriminating phosphate from arsenate is a paramount challenge. Enzymes that utilize phosphate, for example, have the same binding mode and kinetic parameters as arsenate, and the latter's presence therefore decouples metabolism[6,7]. Can proteins discriminate between these two anions, and how would they do so? In particular, cellular phosphate uptake systems face a challenge in arsenate-rich environments. Here we describe a molecular mechanism for this process. We examined the periplasmic phosphate-binding proteins (PBPs) of the ABC-type transport system that mediates phosphate uptake into bacterial cells, including two PBPs from the arsenate-rich Mono Lake *Halomonas* strain GFAJ-1. All PBPs tested are capable of discriminating phosphate over arsenate at least 500-fold. The exception is one of the PBPs of GFAJ-1 that shows roughly 4,500-fold discrimination and its gene is highly expressed under phosphate-limiting conditions. Sub-ångström-resolution structures of *Pseudomonas fluorescens* PBP with both arsenate and phosphate show a unique mode of binding that mediates discrimination. An extensive network of dipole–anion interactions[8,9], and of repulsive interactions, results in the 4% larger arsenate distorting a unique low-barrier hydrogen bond. These features enable the phosphate transport system to bind phosphate selectively over arsenate (at least 10^3 excess) even in highly arsenate-rich environments.

Phosphate transport is mediated by two distinct systems in bacteria and archaea: the phosphate inorganic transport (Pit), a low-affinity and low-selectivity system operating at high phosphate concentrations[10], and the phosphate-specific transport (Pst)[10]. The latter is a high-affinity and high-selectivity ATP-fuelled transporter operating at low phosphate concentration[11] and in high arsenate conditions[10,12]. It comprises a periplasmic high-affinity PBP (or pstS) whose K_d for phosphate is in the submicromolar range[13]. The PBP-captured phosphate anion is subsequently transported across the membrane by an ABC transporter that is energized by ATP to transfer phosphate against the concentration gradient[11]. We examined five bacterial PBPs for their discrimination of phosphate from arsenate. Two PBPs came from arsenate-sensitive species (*Escherichia coli and P. fluorescens*), and the other three came from arsenate-resistant bacteria: *Klebsiella variicola* from Lake Albano (Italy)[14] and the highly arsenate-resistant bacterium *Halomonas* sp. GFAJ-1 (Mono Lake, USA), for which controversy exists over whether it incorporates arsenate into its DNA[5] or not[15,16]. The genome of the latter contains two pstS paralogues from two distinct *pst* operons belonging to so-called 'arsenic islands'[17]. To what degree do these five PBPs distinguish between phosphate and arsenate? What is the molecular basis of their discrimination? Do PBPs adapt to arsenate-rich environments by evolving higher selectivity for phosphate, or perhaps by divergence to import arsenate under certain circumstances?

The five PBPs studied share little sequence identity (Supplementary Fig. 1 and Supplementary Table 1). However, the two *Halomonas* GFAJ-1 PBPs are closely related (Supplementary Fig. 2), and one of them, PBP-2, resides in a phylogenetic clade that comprises extremophiles, and in particular high-salt-resistant and arsenate-resistant organisms (Supplementary Fig. 3 and Supplementary Table 2). We cloned, purified and characterized these five PBPs (Supplementary Information). Despite being isolated from very different sources, four PBPs discriminated phosphate from arsenate 500–850-fold. The exception was the *Halomonas* GFAJ-1 PBP-2 paralogue, which showed much higher discrimination (at least 4,500-fold; Fig. 1a and Supplementary Table 3). Both PBPs of *Halomonas* GFAJ-1 showed specificity for phosphate, which ruled out the possibility that one of them had diverged to serve in arsenate transport. However, the presence of two PBPs raised the question of which of them might be active under phosphate-limiting (and in particular arsenate-rich) environments. Analysis of messenger RNA levels showed that PBP-2 is upregulated more than 40-fold and highly expressed at limiting phosphate concentrations (10 μM), whereas PBP-1 shows barely detectable expression that is only moderately affected by the phosphate level independent of the presence or absence of arsenate (Supplementary Tables 4a, b and 5).

All PBPs have a similar fold that consists of two structurally similar domains linked by a flexible hinge[9] (Supplementary Fig. 4a, b). In the closed form, the binding pocket is buried between the domains and the bound ion is totally dehydrated and completely sequestered. The structure opens to release its cargo anion on binding the transporter. Dipole–ion hydrogen bonds underline the high selectivity of PBPs[18–21] as manifested in the reported 10^5-fold discrimination of phosphate over sulphate[8,9,22]. In both *E. coli* and *P. fluorescens* PBP (the latter dubbed PfluDING), phosphate is bound through 12 hydrogen bonds, of which 9 are ion–dipole interactions: 5 backbone NH groups and 4 hydroxyl groups of serine and threonine (Supplementary Fig. 5).

Phosphate and sulphate differ significantly, not only in having three and two ionizable oxygens, respectively (the structure of sulphate is $(\mathrm{O}{=})_2\mathrm{S}\text{–}\mathrm{O}_2^{2-}$), and in having different pK_a values, but also in size[3]. Indeed, sulphate and similar anions such as iodate are readily distinguished by phosphate-using enzymes[23]. It is less clear, however, how the far more similar phosphate and arsenate are discriminated. In fact, the differences between their binding interactions are so subtle that they could only be unravelled in sub-ångström-resolution structures obtained for the arsenate-bound PfluDING at pH 4.5 and 8.5 (Supplementary Tables 7 and 9; PDB 4F19 and 4F18; 0.95 and 0.96 Å resolution, respectively). The arsenate structures were compared with the equivalent structures with phosphate (Supplementary Tables 8 and 9; PDB 4F1U and 4F1V; 0.98 and 0.88 Å resolution, respectively). Besides the observable difference in the electronic content of phosphorus and arsenic atoms (15 versus 33 electrons), the substitution of phosphate by arsenate was verified by anomalous scattering X-ray

---

[1]Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel. [2]Unité de recherche sur les maladies infectieuses et tropicales émergentes, Faculté de Médecine et de Pharmacie, CNRS-Université de la Méditerranée, 13385 Marseille, France. [3]Institute of Microbiology, Eidgenössische Technische Hochschule Zurich, Wolfgang Pauli Strasse 10, 8093 Zurich, Switzerland.
*These authors contributed equally to this work.

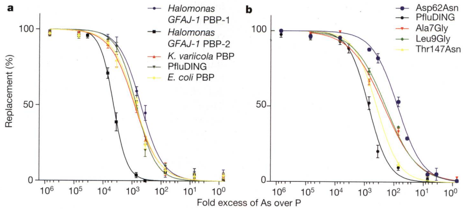


Figure 1 | The phosphate–arsenate selectivity of PBPs. a, PBPs were equilibrated with radiolabelled phosphate and varying arsenate concentrations (0.1 μM to 100 mM). The level of radioactivity (that is, of protein-bound phosphate) with no arsenate corresponded to 0% replacement. *Halomonas* PBP-2 was also found to have higher selectivity for vanadate and sulphate (Supplementary Fig. 12). **b**, The phosphate–arsenate selectivity of PfluDING mutants. The mutation Thr147Asn decreased phosphate binding[18] but hardly affected selectivity. In contrast, mutations in the residues mediating unfavourable interactions (Ala7Gly and Leu9Gly), and mutation Asp62Asn altering the short hydrogen bond, decreased selectivity between fourfold and tenfold (see also Supplementary Figs 9 and 10). Details of the curve fitting and the inferred discrimination factors are provided in Supplementary Tables 3 and 14. Error bars show s.e.m. calculated with two to four independent repeats.

diffraction (Supplementary Table 10 and Supplementary Fig. 6). As observed in arsenate-bound enzyme structures[23,24], the arsenate binds in exactly the same mode as phosphate, and the side chains of the protein within the binding cleft superimpose perfectly (Fig. 2a). Further, the hydrogen locations as determined from the high-quality electronic density maps (Fig. 2b and Supplementary Fig. 7) reveal that, as expected from their identical pK_a values, arsenate and phosphate are both bound in their dibasic forms ($H(As/P)O_4^{2-}$) at both pH 4.5 and pH 8.5 (also confirmed by As–O bond lengths; Supplementary Table 11). This feature—the binding of dibasic phosphate at pH 4.5—was attributed to the unusually high number of hydrogen bonds that solvate the bound anion[22] and to a specific charge network in *Mycobacterium tuberculosis* PBP[21]. Finally, the differences in P–O and As–O bond lengths (0.08–0.14 Å) barely perturb the donor–acceptor hydrogen bond distances (D–H) and angles (Supplementary Tables 11–13). How, then, is arsenate discriminated against?

The only significant difference that we could observe between the arsenate-bound and phosphate-bound structures resided in the remarkably short interaction distance (about 2.50 Å) from the carboxylate of Asp 62 to O2 of the anions, a proposed key interaction for specificity[8,21,25]. This distance is typical of low-barrier hydrogen bonds (LBHBs)[26], and is specifically categorized as a heteromolecular negative-charge-assisted hydrogen bond ((−)CAHB). The shared hydrogen atoms involved in these bonds are rarely observable by diffraction techniques because of their diffuse character[27]. Nonetheless, the high resolution and quality of our structures allowed the shared proton to be located, although at relatively low signal-to-noise ratios (Fig. 3). In the phosphate complex, the hydrogen atom occupies a nearly central position at pH 8.5 (D–H 1.19 Å; Fig. 3a and Supplementary Table 13) but not at pH 4.5 (Supplementary Table 12), as expected for optimal CAHBs. In the arsenate-bound structure, this short donor–acceptor distance is maintained (2.50 Å). However, the

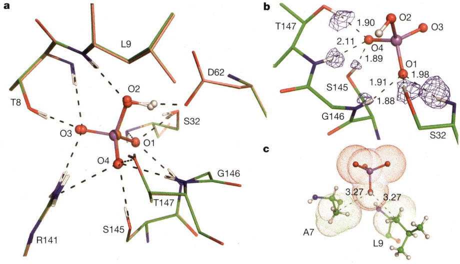


Figure 2 | Binding of arsenate and phosphate to PfluDING. **a**, Superimposition of the arsenate-bound (green sticks; PDB 4F18; pH 8.5) and phosphate-bound (pink sticks; PDB 4F1V; pH 8.5) PfluDING structures. The arsenate (red and violet spheres) and phosphate (orange and red spheres) superimpose perfectly, as do all protein residues. Hydrogen bonds underlining the bound anion are shown as black dashes. **b**, Key ion–dipole interactions in the arsenate structure. Residues involved in the binding (green sticks) donate hydrogen atoms (white sticks, revealed by the omit-H Fourier difference map (blue mesh) contoured at 2σ) to the bound arsenate ion. **c**, Unfavourable contacts between the arsenate ion (red and violet) and the Cβ atoms of Ala 7 and Leu 9 (see also Supplementary Fig. 13). The van der Waals radii are shown as dashes. Distances are indicated in ångströms (standard deviations are available in Supplementary Tables 13 and 11).

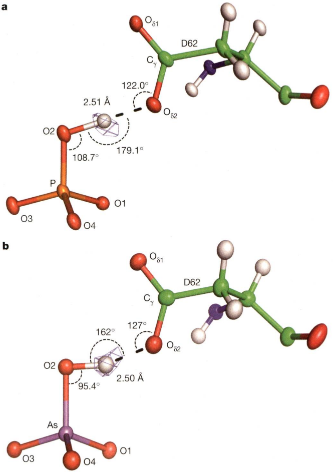


Figure 3 | The (−)CAHB angles are optimal in the phosphate-bound structure but distorted with arsenate. a, A close-up view of the short hydrogen bond ((−)CAHB, or LBHB) between O2 of the bound phosphate and the carboxylate of Asp 62 (pH 8.5; PDB 4F1V). The shared hydrogen atom is revealed by the omit-H Fourier difference map (blue mesh) contoured at 2.5σ (peak height 3.3σ). **b,** The same bond in the arsenate-bound structure, with the omit-H Fourier difference map (blue mesh) contoured at 1.4σ (peak height 1.9σ; the pH 4.5 structure is presented for clarity). Noted are the bond lengths and the angles of the short hydrogen bonds. Atoms are represented as their anisotropic thermal ellipsoids. Measurements made in structures at both pH values (4.5 and 8.5) and corresponding standard deviation values are available in Supplementary Tables 11–13. Alternative models are presented in Supplementary Fig. 8.

proton is asymmetrically located at both pH values (D–H ≈ 1.08 Å), suggesting a weaker interaction (Supplementary Tables 12 and 13).

The three angles that define the CAHB are nearly canonical in the phosphate structure (P–O2–H = 108.7°; O2–H···$O_{\delta2}$ = 179.1°; C_γ-$O_{\delta2}$···H = 122°; which are all ±2° from the canonical angles). However, these angles are all suboptimal for the bound arsenate (As–O2–H = 95.4°; O2–H···$O_{\delta2}$ = 162°; C_γ–$O_{\delta2}$···H = 127; Fig. 3 and Supplementary Tables 11–13). This distortion is the consequence of the longer As–O2 bond than the P–O2 bond (Supplementary Fig. 8b), and may readily account for the difference of about 500-fold in favour of phosphate binding. We note that even when the shared hydrogen was modelled with the canonical As–O2–H angle, and not according to the observed residual density, the remaining two angles were severely distorted (Supplementary Fig. 8a).

The distorted bonding angles with Asp 62 suggest that the primary contribution of Asp 62 is not in anion binding itself. Indeed, the mutation Asp81Asn in *E. coli* PBP (equivalent to Asp62Asn) showed no effect on phosphate binding[25]. However, we found that this mutation, in PfluDING, significantly disturbed the discrimination against arsenate (about tenfold; Fig. 1b and Supplementary Table 14). The equivalent mutation in *E. coli* PBP gave a similar decrease in arsenate discrimination (Supplementary Fig. 9) and an even greater effect with sulphate (about 100-fold; Supplementary Fig. 10). Conversely, a mutation that affected the phosphate-binding ability (Thr147Asn)[18] had little effect on the discrimination (Fig. 1b and Supplementary Table 14). The energy of this short bond may therefore be channelled not towards anion-binding affinity but rather towards anion selectivity.

In that conformational flexibility mediates promiscuity[28], the high selectivity of PBPs relies on structural rigidity. This rigidity relates to the protein's conformation as well as extremely precise positioning and tight packing of the phosphate[21]. The former is illustrated by an analysis of *B*-factors of the region of the binding site (Supplementary Table 15). Moreover, most H-donor groups are backbone NH groups coming from the first turns of α-helices, or from residues with a short side chain (serine or threonine) whose rotameric states are fixed by a network of hydrogen bonds. In addition to the extensive constellation of anion–dipole interactions, the bound anions are also immobilized by tight van der Waals interactions. The O2 of the phosphate is equidistant (3.35 Å) from the Cβ atoms of Ala 7 and Leu 9, thus resulting in two short, unfavourable interactions (Fig. 2c). These distances are smaller than the sum of the van der Waals radii (1.92 Å (radius of CH_2 and CH_3) plus 1.54 Å (radius of OH)[29] = 3.46 Å) and become even shorter (and thus more unfavourable) with arsenate (3.27 Å; values for pH 8.5 structures are given in Supplementary Table 11). Although not previously noted, these interactions seem to be conserved in PBPs (Supplementary Fig. 11). Accordingly, the mutations of either Ala 7 or Leu 9 into glycine that may eliminate these repulsive interactions resulted in a marked decrease in arsenate discrimination (about fourfold; Fig. 1b and Supplementary Table 14).

Taken together, our data indicate that PBPs show exquisite selectivity with respect to arsenate, whereas phosphate-binding proteins tested so far, enzymes and transporters alike, show no or low selectivity[7]. PBPs therefore seem to have evolved a unique mode of binding that is capable of distinguishing between the highly similar phosphate and arsenate. Anion binding is mediated by a dense and rigid network of ion–dipole interactions[9,30]. This network is sensitive to geometric changes. However, the perturbation imposed by the slightly larger arsenate is not equally distributed between all bonds but rather it is channelled almost solely to the short, high-energy hydrogen bond which is most sensitive to both angle and distance perturbations (Fig. 3). The channelling of this unique interaction is the outcome of the extremely tight positioning of the anion, mediated also by unfavourable interactions of the bound anions with two juxtaposed Cβ atoms (Fig. 2c) within a rigid, highly connected binding site (Fig. 2a). Weakening of these interactions (for example by site-directed mutagenesis) significantly decreases the discrimination against arsenate (Fig. 1b); however, at least in the case of the short hydrogen bond, this seems to have little effect on phosphate binding. These structural features therefore support the notion that anion binding and anion selectivity are two independently evolved features.

The notion that PBPs evolved to meet arsenate challenges is also consistent with the observation that the *pst* system seems to be essential for the survival of *E. coli* in arsenate medium[12]. It also seems that the selectivity of bacterial PBPs can be further increased in arsenate-rich environments. The PBP-2 paralogue of *Halomonas* GFAJ-1 showed roughly tenfold higher selectivity than any other bacterial PBP we tested (Fig. 1). PBP-2 also belongs to a phylogenetic clade of high-salt-resistant and arsenate-resistant bacterial species (Supplementary Fig. 3 and Supplementary Table 2). The structural features of this clade and the origin of the high selectivity of PBP-2 remain unknown (Supplementary Fig. 4c, d). However, we note that PBP-2 seems to be functional *in vivo* and is selectively expressed at very low phosphate concentrations (10 µM; Supplementary Tables 4a, b, 5 and 16–18). Finally, the observed discrimination factor for PBP-2, *in vitro* as isolated protein (4,500-fold) is strikingly close to the discrimination levels *in vivo*, by which a 4,000-fold excess of arsenate in the medium yielded only a 2.5-fold excess of intracellular arsenate (Supplementary Table 6). Thus, *Halomonas* GFAJ-1 seems to have evolved to extract phosphate[16] and, owing to the unique discrimination mode of its PBP, can do so at arsenate-to-phosphate ratios that are more than 3,000-fold higher than those observed in Mono Lake.

METHODS SUMMARY

Discrimination assays. The five PBPs in this study were cloned, overexpressed in *E. coli* and purified. PBPs were tested for discrimination of phosphate from different anions (arsenate, vanadate and sulphate) by dialyzing them against 50 mM Tris-HCl buffer pH 8.0 containing a total of 140 nM phosphate (with 1 µCi of radiolabelled phosphate, ^{32}P) and increasing concentrations of the competing anion at tenfold serial dilutions (0.1 µM to 100 mM). All experiments were performed with two to four independent repeats and included a control assay without PBP in the dialysis tube. The background absorption rate by 0.1% BSA was over 1,000-fold less than the PBP signal.

Crystallization and diffraction data collection. PfluDING was extensively dialysed against 50 mM Tris-HCl buffer pH 8 containing 10 mM arsenate and 1 mM $CaCl_2$ (to chelate phosphate). Crystals of the arsenate-bound form were subsequently obtained using hanging-drop vapour diffusion as described previously[22], but using 100 mM arsenate solution as buffer at pH 4.5 or 8.5. Diffraction data collections on PfluDING crystals, including anomalous data at the arsenate edge, were performed at the European Synchrotron Radiation Facility (Grenoble, France) at beamlines ID-29 and ID-23-1. For both data collections in Supplementary Table 7, only 180° out of the 360° collected were used for refinement, to minimize the effect of radiation damage.

A complete description of materials and methods is provided in Supplementary Methods.

Received 6 May; accepted 15 August 2012.

Published online 3 October 2012.

1. Kohn, R. A. & Dunlap, T. F. Calculation of the buffering capacity of bicarbonate in the rumen and *in vitro*. *J. Anim. Sci.* **76,** 1702–1709 (1998).
2. Larsen, E. H. & Hansen, S. H. Separation of arsenic species by ion-pair and ion exchange high performance liquid chromatography. *Microchim. Acta* **109,** 47–51 (1992).
3. Kish, M. M. & Viola, R. E. Oxyanion specificity of aspartate-β-semialdehyde dehydrogenase. *Inorg. Chem.* **38,** 818–820 (1999).
4. Westheimer, F. H. Why nature chose phosphates. *Science* **235,** 1173–1178 (1987).
5. Wolfe-Simon, F. *et al.* A bacterium that can grow by using arsenic instead of phosphorus. *Science* **332,** 1163–1166 (2011).
6. Dixon, H. B. F. The biochemical action of arsonic acids especially as phosphate analogues. *Adv. Inorg. Chem.* **44,** 191–227 (1997).
7. Tawfik, D. S. & Viola, R. E. Arsenate replacing phosphate: alternative life chemistries and ion promiscuity. *Biochemistry* **50,** 1128–1134 (2011).
8. Luecke, H. & Quiocho, F. A. High specificity of a phosphate transport protein determined by hydrogen bonds. *Nature* **347,** 402–406 (1990).
9. Quiocho, F. A. & Ledvina, P. S. Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol. Microbiol.* **20,** 17–25 (1996).
10. Willsky, G. R. & Malamy, M. H. Characterization of two genetically separable inorganic phosphate transport systems in *Escherichia coli*. *J. Bacteriol.* **144,** 356–365 (1980).
11. Rao, N. N. & Torriani, A. Molecular aspects of phosphate transport in *Escherichia coli*. *Mol. Microbiol.* **4,** 1083–1090 (1990).
12. Willsky, G. R. & Malamy, M. H. Effect of arsenate on inorganic phosphate transport in *Escherichia coli*. *J. Bacteriol.* **144,** 366–374 (1980).
13. Quiocho, F. A. Atomic structures of periplasmic binding proteins and the high-affinity active transport systems in bacteria. *Phil. Trans. R. Soc. Lond. B* **326,** 341–351 (1990).
14. Butt, A. & Rehman, A. Isolation of arsenite-oxidizing bacteria from industrial effluents and their potential use in wastewater treatment. *World J. Microbiol. Biotechnol.* **27,** 2435–2441 (2011).
15. Reaves, M. L., Sinha, S., Rabinowitz, J. D., Kruglyak, L. & Redfield, R. J. Absence of detectable arsenate in DNA from arsenate-grown GFAJ-1 cells. *Science* **337,** 470–473 (2012).
16. Erb, T. J., Kiefer, P., Hattendorf, B., Gunther, D. & Vorholt, J. A. GFAJ-1 is an arsenate-resistant, phosphate-dependent organism. *Science* **337,** 467–470 (2012).
17. Kim, E.-H. & Rensing, C. Genome of *Halomonas* strain GFAJ-1, a blueprint for fame or business as usual. *J. Bacteriol.* **194,** 1643–1645 (2012).
18. Ahn, S. *et al.* Structure–function relationships in a bacterial DING protein. *FEBS Lett.* **581,** 3455–3460 (2007).
19. Ledvina, P. S., Yao, N., Choudhary, A. & Quiocho, F. A. Negative electrostatic surface potential of protein sites specific for anionic ligands. *Proc. Natl Acad. Sci. USA* **93,** 6786–6791 (1996).
20. Tanabe, M. *et al.* Structures of OppA and PstS from *Yersinia pestis* indicate variability of interactions with transmembrane domains. *Acta Crystallogr. D* **63,** 1185–1193 (2007).
21. Vyas, N. K., Vyas, M. N. & Quiocho, F. A. Crystal structure of *M.tuberculosis* ABC phosphate transport receptor: specificity and charge compensation dominated by ion–dipole interactions. *Structure* **11,** 765–774 (2003).
22. Liebschner, D. *et al.* Elucidation of the phosphate binding mode of DING proteins revealed by subangstrom X-ray crystallography. *J. Am. Chem. Soc.* **131,** 7879–7886 (2009).
23. Faehnle, C. R., Blanco, J. & Viola, R. E. Structural basis for discrimination between oxyanion substrates or inhibitors in aspartate-β-semialdehyde dehydrogenase. *Acta Crystallogr. D* **60,** 2320–2324 (2004).
24. Kline, P. C. & Schramm, V. L. Purine nucleoside phosphorylase. Catalytic mechanism and transition-state analysis of the arsenolysis reaction. *Biochemistry* **32,** 13212–13219 (1993).
25. Wang, Z., Luecke, H., Yao, N. & Quiocho, F. A. A low energy short hydrogen bond in very high resolution structures of protein receptor–phosphate complexes. *Nature Struct. Biol.* **4,** 519–522 (1997).
26. Cleland, W. W. Low-barrier hydrogen bonds and enzymatic catalysis. *Arch. Biochem. Biophys.* **382,** 1–5 (2000).
27. Steiner, T. & Saenger, W. Lengthening of the covalent O–H bond in O–H···O hydrogen bonds re-examined from low-temperature neutron diffraction data of organic compounds. *Acta Crystallogr. B* **50,** 348–357 (1994).
28. Khersonsky, O. & Tawfik, D. S. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **79,** 471–505 (2010).
29. Li, A. J. & Nussinov, R. A set of van der Waals and coulombic radii of protein atoms for molecular and solvent-accessible surface calculation, packing evaluation, and docking. *Proteins* **32,** 111–127 (1998).
30. Pflugrath, J. W. & Quiocho, F. A. Sulphate sequestered in the sulphate-binding protein of *Salmonella typhimurium* is bound solely by hydrogen bonds. *Nature* **314,** 257–260 (1985).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank G. Gotthard for valuable inputs, C. Jelsch for help with the software Mopro, and A. Toth-Petroczy for help with the phylogenetic analysis. Financial support by the Israel Science Foundation is gratefully acknowledged. M.E. is a fellow supported by the Intra-European Fellowships Marie Curie program (grant No. 252836). D.S.T. is the Nella and Leon Benoziyo Professor of Biochemistry. T.J.E. was supported by an Eidgenössische Technische Hochschule fellowship.

Author Contributions M.E. initiated the project, designed experiments, crystallized the proteins, collected and processed crystallographic data, performed discrimination assays, analysed the data and wrote the manuscript. A.W. cloned, purified and crystallized proteins, performed discrimination assays and analysed data. K.G.A. cloned and constructed the mutants, purified proteins and performed discrimination assays. T.J.E. and J.A.V. designed, performed and analysed reverse transcription PCR, as well as cellular experiments, and provided physiological data. E.C. purified proteins and collected diffraction data. D.S.T. initiated the project, designed experiments, analysed data and wrote the manuscript.

Author Information Coordinates and structure factors for the arsenate-bound PfluDING structures at pH 4.5 and 8.5 are deposited in the Protein Data Bank (accession codes 4F18 and 4F19); the phosphate-bound PfluDING coordinates and structures at pH 4.5 and 8.5 (re-refinement) are deposited in the Protein Data Bank (accession codes 4F1U and 4F1V). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.S.T. (tawfik@weizmann.ac.il) or M.E. (mikael.elias@weizmann.ac.il).

doi:10.1038/nature11500

The spatial architecture of protein function and adaptation

Richard N. McLaughlin Jr[1,2]†, Frank J. Poelwijk[1,2], Arjun Raman[1,2], Walraj S. Gosal[1,2] & Rama Ranganathan[1,2]

Statistical analysis of protein evolution suggests a design for natural proteins in which sparse networks of coevolving amino acids (termed sectors) comprise the essence of three-dimensional structure and function[1–5]. However, proteins are also subject to pressures deriving from the dynamics of the evolutionary process itself—the ability to tolerate mutation and to be adaptive to changing selection pressures[6–10]. To understand the relationship of the sector architecture to these properties, we developed a high-throughput quantitative method for a comprehensive single-mutation study in which every position is substituted individually to every other amino acid. Using a PDZ domain (PSD95$^{\text{pdz3}}$) model system, we show that sector positions are functionally sensitive to mutation, whereas non-sector positions are more tolerant to substitution. In addition, we find that adaptation to a new binding specificity initiates exclusively through variation within sector residues. A combination of just two sector mutations located near and away from the ligand-binding site suffices to switch the binding specificity of PSD95$^{\text{pdz3}}$ quantitatively towards a class-switching ligand. The localization of functional constraint and adaptive variation within the sector has important implications for understanding and engineering proteins.

A basic tenet of biology is that the amino acid sequence of proteins specifies their three-dimensional structure and biochemical function, at least in a physiological setting[11]. Statistical coupling analysis (SCA)[1,2,5] is a quantitative approach for understanding the information content of protein sequences through a generalization of the principle of evolutionary conservation. The underlying premise is that the pattern of energetic couplings between residues in a protein—the functional constraints between amino acids—might be exposed through a statistical analysis of coevolution of those residue positions in a family of homologous sequences. A main conclusion of SCA is that most residues in proteins evolve nearly independently, without much influence from even their immediate structural environment, whereas about 20% of amino acids (Fig. 1a and Supplementary Fig. 1) are organized into physically contiguous networks of coevolving amino acids, termed protein sectors[1,2]. Sectors are typically built around protein active sites, but connect to distant functional surfaces through pathways of residue interactions in the protein core[12]. For example, in the PDZ family of protein interaction modules, the sector connects the ligand-binding pocket with an allosteric site (Fig. 1b, asterisk) on the opposite surface[2,13]. Sectors are found in every protein family studied so far and are related to conserved functional activities, suggesting that this structural feature is a general property of natural proteins[1,3,5,14–16].

An important next step is to understand why the design of natural proteins should look like the sector architecture, with distributed sparse networks of cooperatively acting residues embedded within an environment of weakly coupled amino acids. One reason might be necessity for native folding and function, but this seems unlikely. The marked recent advances in physics-based protein design are largely based on homogeneous optimization of local interactions in protein structures[17,18], and alternative explanations for long-range communication within proteins have been suggested[19]. Our proposal is that the sector is the natural consequence of evolutionary constraints not typically considered in protein engineering or biophysical models, primarily of the need for adaptive variation in response to fluctuating conditions of fitness. By placing the constraints on native folding and function on sector positions, this architecture might provide the capacity for rapid adaptive variation through mutation of a few cooperatively acting residues. If so, the plurality of non-sector positions, regardless of structural location, should display much more mutation tolerance and less adaptive potential.

Previous studies have tested the role of sectors in protein function using targeted mutagenesis of a few amino acid positions[1–3,13–16,20]. Although useful, these studies cannot convincingly test the sector hypothesis posed here, mainly due to the limited scale of experimentation. To address this, we developed a quantitative high-throughput method based on next-generation sequencing suitable for a large-scale mutational analysis of proteins in a cellular context (Fig. 2 and Supplementary Fig. 2). The method is implemented here for comprehensive single mutagenesis in one representative member of the PDZ family of protein interaction modules (PSD95$^{\text{pdz3}}$) as a model system, but could be used for the study of many other proteins and for higher-order mutational studies (F.J.P. and R.R., unpublished observations).

The method involves three components: (1) a bacterial two-hybrid (B2H) system modified from a previous study[21] in which the ability of PSD95$^{\text{pdz3}}$ to bind its cognate ligand (-TKNYKQTSV-COOH, derived from the cysteine-rich interactor of PDZ (CRIPT)[22]) is quantitatively linked to the expression of enhanced green fluorescent protein (eGFP) (Supplementary Fig. 2a); (2) a fluorescence-activated cell sorting (FACS) step, in which bacterial populations carrying large libraries of mutations in the protein are selected for those cells displaying

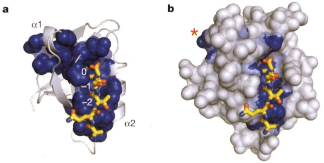


Figure 1 | Sector architecture in the PDZ domain family. a, **b**, The PDZ sector (blue spheres) shown in a cartoon (**a**) or space filling (**b**) representation of the structure of rat PSD95$^{\text{pdz3}}$ (Protein Data Bank (PDB) accession 1BE9). Yellow stick bonds represent the co-crystallized peptide ligand, with ligand positions numbered (0, −1, −2). The sector comprises a sparse network of residues built around the ligand-binding pocket and connecting to a distant surface site (marked with asterisk) through a subset of amino acid interactions within the protein core.

---

[1]Green Center for Systems Biology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9050, USA. [2]Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9050, USA. †Present address: Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Seattle, Washington 98109-1024, USA.

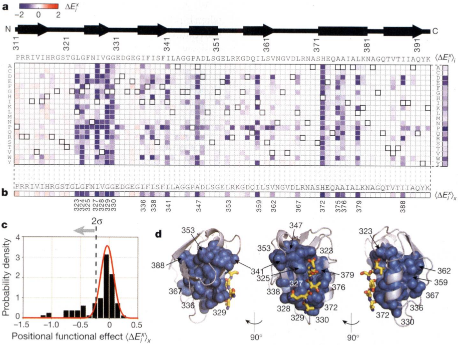


Figure 2 | Complete single mutagenesis in PSD95$^{\mathrm{pdz3}}$. **a**, The data matrix showing ΔE_i^x—the functional cost of every mutation x at each position i relative to wild-type PSD95$^{\mathrm{pdz3}}$—colorimetrically, with blue representing loss-of-function and red representing gain-of-function mutations. The wild-type amino acid at each position is indicated by bold squares in the grid. The average functional cost of each amino acid substitution over all positions ($\langle \Delta E_i^x \rangle_i$) is shown at the right. **b**, The functional cost of all amino acid substitutions at each position shown as the average taken over each column ($\langle \Delta E_i^x \rangle_x$). **c**, A histogram of the data in panel **b** indicates positions with a significant effect ($>2\sigma$, 20 out of 83). **d**, Mapping of the 20 functionally significant positions on the PSD95$^{\mathrm{pdz3}}$ structure, the peptide ligand shown as yellow stick bonds. These positions comprise a distributed, physically contiguous network built around the binding pocket and extending through the protein structure.

eGFP levels above a specified threshold (Supplementary Fig. 2b, c); and (3) Solexa high-throughput sequencing to determine the frequency of each allele in the unselected and selected populations [23–26] (Supplementary Fig. 2c). The effect of each mutation is then expressed as the log frequency of observing each amino acid x at each position i in the selected (sel) versus the unselected (unsel) population, relative to wild type (WT):

$$\Delta E_i^x = \log\left[\frac{f_i^{x,\mathrm{sel}}}{f_i^{x,\mathrm{unsel}}}\right] - \log\left[\frac{f_i^{\mathrm{WT,sel}}}{f_i^{\mathrm{WT,unsel}}}\right]$$

In this assay, mutations that show no functional effect should show a relative frequency in the selected population that is identical to wild type ($\Delta E_i^x \approx 0$), and deviations from this expectation should provide a quantitative measure of the functional effect of each mutation. Tuning of growth and induction parameters and introduction of a point mutation in the bateriophage λ-cI at Glu 34 (ref. 27) (Supplementary Fig. 3) led to experimental conditions that showed a near-linear relationship between the binding free energy of the PSD95$^{\mathrm{pdz3}}$ ligand interaction (ΔG_{bind}) and ΔE_i^x over the range of binding affinities reported for natural PDZ domains (~0.1–200 μM; Supplementary Figs 2d and 4).

We used the B2H–sequencing assay to carry out a complete single-mutation scan for the PSD95$^{\mathrm{pdz3}}$ domain in which every position shared with the overall PDZ family (83 total) is individually mutated to every other amino acid (Fig. 2a, 83 positions × 19 mutations + wild type = 1,578 variants). The data reveal several aspects of the global pattern of mutational sensitivity in the protein. First, the effect of each amino acid substitution averaged over all positions (across rows in the data matrix) is what might be predicted from the chemical properties of the side chains ($\langle \Delta E_i^x \rangle_i$, Fig. 2a, right). Proline is the most unfavourable substitution, followed by amino acids that are formally charged at neutral pH (Asp, Glu, Lys and Arg), and by tryptophan, the volumetrically largest side chain. Substitutions to alanine or cysteine introduce the least perturbation on average, consistent with their use for scanning mutation or solvent accessibility studies, respectively.

To examine the position-specific effects of mutation in PSD95$^{\mathrm{pdz3}}$, we considered both the full data matrix (Fig. 2a) and the average effect of all mutations per position $\langle \Delta E_i^x \rangle_x$ (Fig. 2b, c). This analysis indicates a heterogeneous, distributed and physically contiguous network of functional residues in PSD95$^{\mathrm{pdz3}}$ (Fig. 2b–d). Most positions show little effect on mutation, tolerating nearly every substitution even if radically different in chemical character (Fig. 2a–c and Supplementary Fig. 5). This includes some that are in direct contact with peptide ligand (for example, 326 and 380), and some that are buried in the protein core and largely conserved (for example, 314, 316, 356, 357 and 390). By contrast, a subset of positions (20 out of 83, Fig. 2c) shows significant sensitivity to mutation ($>2\sigma$ from mean). Within the binding pocket, His 372 tolerates essentially no other substitution and Leu 323, Phe 325, Ile 327, and Leu 379 show tolerance to only the most chemically conservative mutations. However, outside the direct spatial environment of the ligand, Gly 329, Gly 330, Ile 336, Ala 347, Leu 353, Val 362 and Ala 375 comprise a subset of buried residues that also show significant sensitivity to mutation—the largest average mutational effect in the whole protein comes from position 329.

Analysis of the relationship between this global mutational analysis and the protein sector in the PDZ domain (Fig. 3a, top) shows that sector positions selectively comprise the tail region of the distribution of mutational sensitivity. Of 81 positions tested and statistically well-represented in the PDZ multiple sequence alignment (MSA), 20 positions show a significant functional effect (Fig. 2b, c), and 15 of these are sector positions (out of 20 sector positions in total) (Fig. 1 and Supplementary Fig. 1), indicating very strong statistical correlation between sector positions and functional effect on mutation ($P < 10^{-8}$, Fisher's exact test, Supplementary Fig. 6a). This correlation is robust to cut-offs used for both categories (Supplementary Fig. 6b).

Furthermore, the analysis shows that more standard measures for predicting functional importance of amino acids—such as burial in the

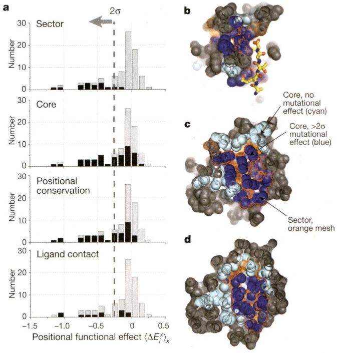


Figure 3 | **The relationship of mutational sensitivity of positions to the protein sector.** **a**, The distribution of mutational effect in PSD95$^{\mathrm{pdz3}}$ (grey), overlaid with distributions (in black) of sector, core (solvent accessibility < 0.15), positionally conserved (relative entropy >1, the mean value over all positions in the MSA) and ligand contacting positions (within 4 Å shell of ligand atoms). **b**–**d**, Slices through the core of PSD95$^{\mathrm{pdz3}}$, showing mutationally significant core positions (dark blue), mutationally non-significant core positions (cyan), and the sector (orange mesh). All non-core positions are in grey, with the peptide ligand shown in yellow stick bonds.

protein core, positional conservation and contact with ligand—can also identify functional sites, but these criteria are not good overall descriptors of the data. Indeed, there are more conserved and buried positions that show no significant effect on mutation than otherwise (Fig. 3a, middle panels). Spatial proximity to ligand is the least likely hypothesis for explaining functional importance (Fig. 3a, bottom), consistent with the observations that not every direct interaction with ligand contributes to binding energy[28] and that non-local sites can influence active site function indirectly[2]. Serial slices through the core of the PSD95$^{\mathrm{pdz3}}$ domain are also consistent with these observations; the sector (Fig. 3b–d, orange mesh) largely captures the functional subset of core positions (dark blue spheres) embedded among many other non-functional core positions (cyan spheres). We conclude that the main functional constraints on the PDZ domain are localized within the distributed network of amino acid positions that define the sector.

To determine whether the sector specifically encodes the ability to adapt to altered selection pressures, we repeated the global single-mutation study in PSD95$^{\mathrm{pdz3}}$, challenging the domain to bind a non-native peptide ligand (T$_{-2}$F, TKNYKQ_F_SV-COOH, indicating a Thr to Phe mutation at the minus two position (Fig. 1a); ligand positions are numbered in reverse order from the carboxy terminus (position 0)) (Supplementary Fig. 7). T$_{-2}$F switches the CRIPT PDZ ligand from class I (-S/T-X-Ψ-COOH, where X is any amino acid, Ψ is hydrophobic) to class II (-X-Ψ-X-Ψ-COOH) specificity, and represents a substantial but physiologically relevant variation in PDZ function[29]. Accordingly, PSD95$^{\mathrm{pdz3}}$ binds to the T$_{-2}$F peptide with an approximately 45-fold decrease in binding affinity compared with the wild-type class I ligand (Fig. 4f).

We measured the difference in the functional effect of mutations when binding either the wild-type or T$_{-2}$F ligands, a global analysis of the context dependence (or epistasis) between every mutation of every position in PSD95$^{\mathrm{pdz3}}$ and the T$_{-2}$F mutation on the peptide ligand (Fig. 4a, b). The data show that nearly all mutations at positions in PSD95$^{\mathrm{pdz3}}$—even if they have an absolute effect on ligand binding—show the same effect on binding the T$_{-2}$F ligand as for the wild-type ligand, resulting in no significant epistasis with the T$_{-2}$F mutation (Fig. 4c). These positions are insensitive to the switch in peptide identity and would therefore be non-adaptive with regard to this perturbation. However, mutations at nine positions in PSD95$^{\mathrm{pdz3}}$ show statistically significant epistasis with the T$_{-2}$F mutation (Fig. 4b, c); these positions show a mutational response that depends on the identity of the target ligand. These positions uniformly show an effect on mutation that is less deleterious for the T$_{-2}$F ligand than for the wild-type ligand (Fig. 4b, red pixels). All of these positions are within the sector (Fig. 4d, see overlap between sector (blue mesh) and epistatic positions (red)), and map out a pattern of epistasis for the T$_{-2}$F mutation that involves a spatially distributed network in the PDZ protein structure comprising residues both near and far from the minus two ligand position (Figs. 1a and 4d).

A subset of the sector positions (322, 330, 336 and 372) shows such extreme epistasis that the positions flip in the direction of mutational effect between the wild-type and T$_{-2}$F ligands (Fig. 4a, b, d). Mutations at these positions destabilize or are neutral for binding to the wild-type ligand, but are favourable for binding the T$_{-2}$F ligand and would therefore be selected when challenged to bind this class-switching variant. Notably, only one of the selectable sites (372) directly contacts the T$_{-2}$ ligand position. The other sites are located either one shell (329, 330) or two shells (336) of residues away from the binding pocket, or act at a distance from a loop contacting the terminal carboxylate of the ligand (322; Fig. 4d). We also find that for these few sites of selection for T$_{-2}$F, it is not merely one or two mutations, but nearly every substitution that shows the effect of stabilizing binding to the T$_{-2}$F ligand (Fig. 4b and Supplementary Fig. 7).

These data describe many potential mutational routes for switching the binding specificity PSD95$^{\mathrm{pdz3}}$ towards the class II T$_{-2}$F ligand. For example, mutation of position 330, which does not contact ligand (Fig. 4d), to threonine is expected to moderately destabilize binding to wild-type CRIPT ligand, but is the most favourable of all single mutants for the T$_{-2}$F ligand (Fig. 4e). Also, His372Ala, a mutation at a position directly linking the -2 position of the ligand and position 330 (Fig. 4d), is expected to decrease binding for CRIPT more strongly than the Gly330Thr mutation, but to increase binding for T$_{-2}$F to a similar degree (Fig. 4e). Binding affinities for Gly330Thr, His372Ala and the double mutant combination were measured using a fluorescence polarization assay and show excellent consistency with the B2H data (Fig. 4f). His372Ala converts PSD95$^{\mathrm{pdz3}}$ from a protein with an approximately 45-fold preference for the CRIPT ligand to one with about a 14-fold preference for the T$_{-2}$F ligand, a partial specificity switch from a single mutation. By contrast, Gly330Thr converts PSD95$^{\mathrm{pdz3}}$ to a domain with an unexpected phenotype: high-affinity but non-specific recognition of both CRIPT and T$_{-2}$F ligands. Such a phenotype could be evolutionarily important when a mutational path characterized by a promiscuous but biologically functional intermediate is advantageous. Finally, the combination of both mutations (Gly330Thr and His372Ala) completes the specificity switch; this double mutant displays an approximately 45-fold preference for the T$_{-2}$F ligand. These data demonstrate that short, cooperative paths of mutation within the sector can suffice to change functional specificity quantitatively.

In summary, the data presented here indicate that protein robustness and adaptability can be explained through a model in which the main functional constraints are loaded in the sector—a sparse, collectively evolving network within the protein structure. By saturation point mutagenesis, we find that sector positions selectively resist variation when challenged with wild-type ligand, but can flip to promote

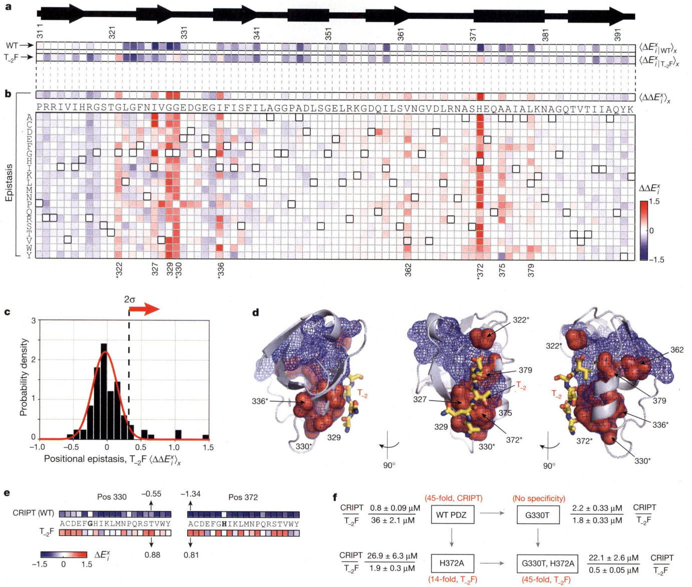


Figure 4 | Adaptation through sector variation. **a**, Average mutational effect in PSD95$^{\mathrm{pdz3}}$ when binding the wild-type ligand CRIPT ($\left\langle \Delta E^x_{i|\mathrm{WT}} \right\rangle_x$, top) or a class-switching $T_{-2}F$ variant ($\left\langle \Delta E^x_{i|T_{-2}F} \right\rangle_x$, bottom). $T_{-2}F$ contains a Thr to Phe mutation at position minus two of the peptide ligand (Fig. 1a). **b**, The difference, or epistasis, between mutational effects for binding wild-type or $T_{-2}F$ ligands, shown either averaged over amino acids at each position (top, $\langle \Delta\Delta E^x_i \rangle_x$) or broken down by amino acid (bottom, $\Delta\Delta E^x_i = \Delta E^x_{i|T_{-2}F} - \Delta E^x_{i|WT}$). The nine positions showing statistically significant epistasis (**c**) are numbered, and the asterisks mark positions where mutations on average can be positively selected for the $T_{-2}F$ ligand. **c**, A histogram of epistasis between mutations at each position in PSD95$^{\mathrm{pdz3}}$ and the $T_{-2}F$ ligand variation. **d**, A mapping of epistatic positions (red) on the structure of PSD95$^{\mathrm{pdz3}}$; the wild-type peptide ligand is shown in yellow stick bonds, the T_{-2} position is indicated in red and sector positions are in blue mesh. The positions showing mutational epistasis with $T_{-2}F$ comprise a physically distributed network propagating from the T_{-2} position, and are entirely composed of sector positions. The asterisks are as in **b**. **e**, The B2H–sequencing data for sector positions 330 and 372 when binding CRIPT (Fig. 2c) or $T_{-2}F$ (Supplementary Fig. 7) ligands suggest mutations for altering the specificity of PSD95$^{\mathrm{pdz3}}$ towards $T_{-2}F$; the wild-type amino acid is shown in bold. **f**, Binding affinities for purified PSD95$^{\mathrm{pdz3}}$ carrying the Gly330Thr and His372Ala mutations both singly and together. Gly330Thr displays high affinity but non-specific binding for CRIPT and $T_{-2}F$ ligands, His372Ala shows a partial specificity switch towards $T_{-2}F$ and the double mutant represents a complete specificity switch for these two ligands.

variation when challenged with a functionally distinct ligand. This epistatic coupling between ligand and sector underlies efficient functional adaptation, permitting considerable changes in specificity through very few mutations. Turned around, these data provide support for the hypothesis that the sector architecture might be the natural solution to design by evolution under conditions of constantly fluctuating environments. Such environments impose the need for maintaining robustness to mutation and adaptability to varying selection pressures and have been shown theoretically to influence the design of evolving systems[30]. It will be important to now experimentally test the notion that the statistical history of fluctuations in conditions of selection fundamentally defines the physical design of natural proteins.

METHODS SUMMARY

SCA and sector identification were carried out using version 5.0 of the SCA MATLAB toolbox. The software and a script for carrying out the calculations is available for download from our laboratory website (http://systems.swmed.edu/rr_lab).

Comprehensive single-mutation libraries in PSD95$^{\mathrm{pdz3}}$ were constructed by oligonucleotide-directed mutagenesis, randomizing each codon to NNS, in which N represents a mixture of all four bases, and S represents a mixture of G and C. NNS libraries for each codon were separately amplified by PCR, mixed in an equimolar ratio, and cloned into pZS22-PDZ3 (see later) to make the library.

The B2H system consists of three plasmids: (1) pZS22-PDZ3 (kanamycin resistant), providing IPTG-inducible expression of PSD95$^{\mathrm{pdz3}}$ variants fused to the DNA-binding domain of bacteriophage λ-cI; (2) pZA31-RNAα-CRIPT (or $T_{-2}F$) (chloramphenicol resistant), providing anhydrotetracycline-inducible expression of the amino-terminal domain of *Escherichia coli* RNA polymerase α-subunit fused to the target peptide ligand; and (3) pZE1RM-eGFP (ampicillin resistant), containing the target promoter driving the enhanced green fluorescent protein (eGFP). MC4100-Z1 *E. coli* cells transformed with these three plasmids were grown in ZYM-505 media, diluted with inducers to attenuance at 600 nm ($D_{600\,\mathrm{nm}}$) of 0.4, grown for 2 h, and subjected to FACS sorting (BD FACSAria, gates set at top 10% and 25% of eGFP distribution for wild-type $\mathrm{PSD95}^{\mathrm{pdz3}}$). Sorted cells were grown in ZYM-505 for 12 h, miniprepped and amplified by PCR to prepare samples for Solexa paired-end sequencing completed using Solexa v4 PE-flowcell (University of Texas Southwestern genomics core) and analysed through CLC Genomics Workbench and self-coded software.

For biophysical measurements, $\mathrm{PSD95}^{\mathrm{pdz3}}$ variants were expressed in BL21(DE3) *E. coli* cells as glutathione *S*-transferase (GST) fusions and purified using affinity chromatography and cleavage of the GST tag. Binding affinities were determined using tetramethylrhodamine (TMR)-labelled target peptides, monitoring fluorescence polarization of TMR on a $\mathrm{Victor}^3\mathrm{V}$ plate reader as a function of $\mathrm{PSD95}^{\mathrm{pdz3}}$ concentration.

Full Methods and any associated references are available in the online version of the paper.

Received 18 July 2011; accepted 10 August 2012.

Published online 7 October 2012.

1. Halabi, N., Rivoire, O., Leibler, S. & Ranganathan, R. Protein sectors: evolutionary units of three-dimensional structure. *Cell* **138,** 774–786 (2009).
2. Lockless, S. W. & Ranganathan, R. Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* **286,** 295–299 (1999).
3. Russ, W. P., Lowery, D. M., Mishra, P., Yaffe, M. B. & Ranganathan, R. Natural-like function in artificial WW domains. *Nature* **437,** 579–583 (2005).
4. Socolich, M. *et al.* Evolutionary information for specifying a protein fold. *Nature* **437,** 512–518 (2005).
5. Süel, G. M., Lockless, S. W., Wall, M. A. & Ranganathan, R. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nature Struct. Mol. Biol.* **10,** 59–69 (2003).
6. Bershtein, S., Segal, M., Bekerman, R., Tokuriki, N. & Tawfik, D. S. Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature* **444,** 929–932 (2006).
7. Bloom, J. D., Labthavikul, S. T., Otey, C. R. & Arnold, F. H. Protein stability promotes evolvability. *Proc. Natl Acad. Sci. USA* **103,** 5869–5874 (2006).
8. Draghi, J. A., Parsons, T. L., Wagner, G. P. & Plotkin, J. B. Mutational robustness can facilitate adaptation. *Nature* **463,** 353–355 (2010).
9. Tiana, G., Shakhnovich, B. E., Dokholyan, N. V. & Shakhnovich, E. I. Imprint of evolution on protein structures. *Proc. Natl Acad. Sci. USA* **101,** 2846–2851 (2004).
10. Voigt, C. A., Kauffman, S. & Wang, Z. G. Rational evolutionary design: the theory of *in vitro* protein evolution. *Adv. Protein Chem.* **55,** 79–160 (2000).
11. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181,** 223–230 (1973).
12. Reynolds, K. A., McLaughlin, R. N. & Ranganathan, R. Hot spots for allosteric regulation on protein surfaces. *Cell* **147,** 1564–1575 (2011).
13. Peterson, F. C., Penkert, R. R., Volkman, B. F. & Prehoda, K. E. Cdc42 regulates the Par-6 PDZ domain through an allosteric CRIB-PDZ transition. *Mol. Cell* **13,** 665–676 (2004).
14. Hatley, M. E., Lockless, S. W., Gibson, S. K., Gilman, A. G. & Ranganathan, R. Allosteric determinants in guanine nucleotide-binding proteins. *Proc. Natl Acad. Sci. USA* **100,** 14445–14450 (2003).
15. Shulman, A. I., Larson, C., Mangelsdorf, D. J. & Ranganathan, R. Structural determinants of allosteric ligand activation in RXR heterodimers. *Cell* **116,** 417–429 (2004).
16. Smock, R. R. O., Russ, W. P., Swain, J. F., Leibler, S., Ranganathan, R. & Gierasch, L. M. An interdomain sector mediating allostery in Hsp70 molecular chaperones. *Mol. Syst. Biol.* **6,** 414 (2010).
17. Bradley, P., Misura, K. M. S. & Baker, D. Toward high-resolution *de novo* structure prediction for small proteins. *Science* **309,** 1868–1871 (2005).
18. Kuhlman, B. *et al.* Design of a novel globular protein fold with atomic-level accuracy. *Science* **302,** 1364–1368 (2003).
19. Hilser, V. J. & Thompson, E. B. Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc. Natl Acad. Sci. USA* **104,** 8311–8315 (2007).
20. Chi, C. N. *et al.* Reassessing a sparse energetic network within a single protein domain. *Proc. Natl Acad. Sci. USA* **105,** 4679–4684 (2008).
21. Dove, S. L., Joung J. K. & Hochschild, A. Activation of prokaryotic transcription through arbitrary protein-protein contacts. *Nature* **386,** 627–630 (1997).
22. Niethammer, M. *et al.* CRIPT, a novel postsynaptic protein that binds to the third PDZ domain of PSD-95/SAP90. *Neuron* **20,** 693–707 (1998).
23. Adkar, B. V. *et al.* Protein model discrimination using mutational sensitivity derived from deep sequencing. *Structure* **20,** 371–381 (2012).
24. Fowler, D. M. *et al.* High-resolution mapping of protein sequence-function relationships. *Nature Methods* **7,** 741–746 (2010).
25. Kinney, J. B., Murugan, A., Callan, C. G. & Cox, E. C. Using deep sequencing to characterize the biophysical mechanism of a transcriptional regulatory sequence. *Proc. Natl Acad. Sci. USA* **107,** 9158–9163 (2010).
26. van Opijnen, T., Bodi, K. L. & Camilli, A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature Methods* **6,** 767–772 (2009).
27. Whipple, F. W., Ptashne, M. & Hochschild, A. The activation defect of a λ-cI positive control mutant. *J. Mol. Biol.* **265,** 261–265 (1997).
28. Clackson, T. & Wells J. A. A hot spot of binding energy in a hormone-receptor interface. *Science* **267,** 383–386 (1995).
29. Songyang, Z. *et al.* Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science* **275,** 73–77 (1997).
30. Kussell, E. & Leibler, S. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309,** 2075–2078 (2005).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank W. Russ, P. Mishra and other members of the Ranganathan laboratory for contributions to this work, W. Wakeland and C. Liang for assistance with Solexa sequencing, E. Curry and A. Mobley for assistance with flow cytometry, and M. Elowitz for providing the pZ plasmids. We acknowledge support from the University of Texas Southwestern Graduate School and Pharmacology Training Grant (R. N. M.), the Helen Hay Whitney Fellowship program (F.J.P.) and support from the National Institutes of Health (R01EY018720-05, R.R.), The Robert A. Welch Foundation (I-1366, R.R.) and the Green Center for Systems Biology (R.R.).

Author Contributions R.N.M. and R.R. developed the research plan and experimental strategy. R.N.M. built the B2H assay, collected the data and wrote and executed the code for analysis of the B2H and sequencing data. F.J.P. and W.S.G. improved the dynamic range of the B2H assay. A.R. carried out the mutational analysis in Fig. 4 e, f. R.N.M. and R.R. analysed the data and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.R. (rama.ranganathan@utsouthwestern.edu).

PROMOTION

Make Your Holiday Brighter!

Got a new iPad? Get the new SCIENTIFIC AMERICAN Tablet Edition for iPad.

How Feedbacks Could Speed Warming
Growing an Eye in the Lab
SCIENTIFIC AMERICAN
The Inner Life of Quarks
Science aims to look inside the smallest bits of matter

Get closer to the leading edge of science and technology: experience the SCIENTIFIC AMERICAN Tablet Edition - now available for iPad®.

Available on the App Store

TO RENEW
SCIENTIFICAMERICAN.COM/UPGRADE

TO SUBSCRIBE
SCIENTIFICAMERICAN.COM/TABLET

Apple and iPad are trademarks of Apple Inc., registered in the U.S. and other countries. App Store is a service mark of Apple Inc.

Flow Cytometry
by
Stratedigm
Cytometry Without Compromise

S1000Ex -
Up to 4 lasers, 14 colors

S1000 -
Up to 4 lasers, 8 colors

SE520 -
2 lasers, 4 colors

A600 -
High throughput autosampler

Industry leading sensitivity in a compact footprint.

Stratedigm
Cytometry Without Compromise

www.stratedigm.com
408.512.3910 info@stratedigm.com

ACTIVE MOTIF®
Enabling Epigenetics Research

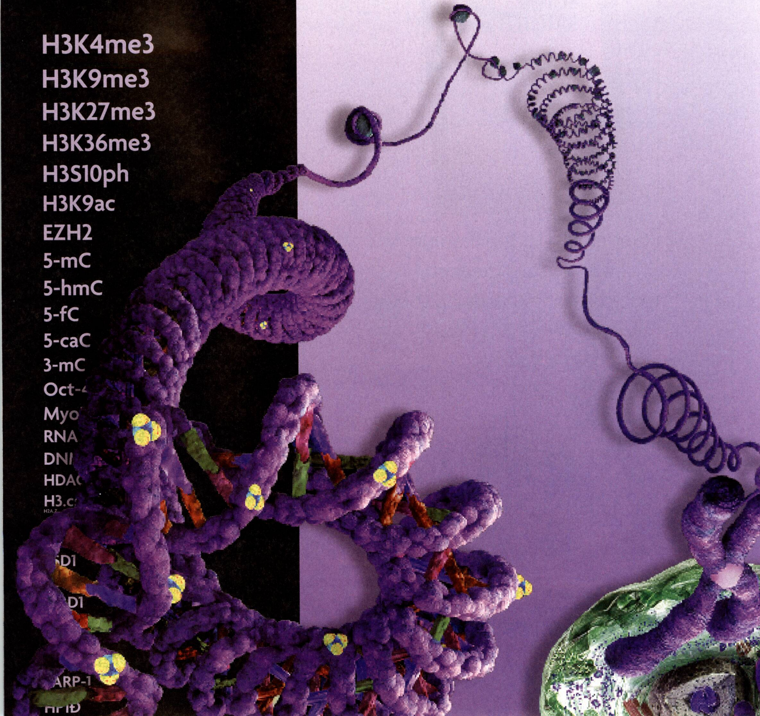
Active Motif is dedicated to developing quality antibodies and kits to characterize key epigenetic events.

Visit us at www.activemotif.com for a complete list of antibodies for DNA methylation , ChIP and ChIP-Seq.

Now offering ChIP-Seq validated antibodies

Superior Antibodies & Kits for
EPIGENETICS RESEARCH

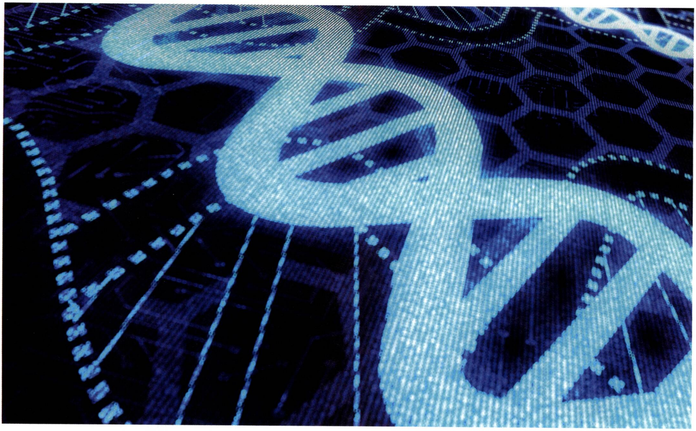
Visit us at www.activemotif.com


H3K4me3
H3K9me3
H3K27me3
H3K36me3
H3S10ph
H3K9ac
EZH2
5-mC
5-hmC
5-fC
5-caC
3-mC
Oct-4
Myo
RNA
DN
HDA
H3.c
H2A.Z
SD1
D1
ARP-1

TECHNOLOGY FEATURE

READING THE SECOND GENOMIC CODE

Transient changes to the genome make its code more complex to interpret but they still put a gleam in the eye of drug and technology developers.


ALENGO/ISTOCKPHOTO

BY VIVIEN MARX

DNA is famous as the instruction manual of life — the multi-billion-base-pair data tape that directs how a fertilized egg turns into the specific cells, tissues and organs of, say, a sharp-eyed soccer pro who is musically inclined but who also battles depression.

But DNA works with many partners, including 'epigenetic' factors, which influence gene expression in ways that don't involve changes to the underlying sequence (see 'Polygamous DNA'). An important example is methylation, in which methyl groups are tacked on to various locations along the double helix to control the activity of particular genes. Methylation also affects histones, the spool-like proteins around which DNA is tightly wound inside the nucleus: the chemical modifications help to control when this protein–DNA complex, called chromatin, opens up so that the genetic instructions can be read.

Figuring out when and how such epigenetic changes get made — or damaged — has become a crucial part of scientists' efforts to understand both the normal development of cells and their progression into cancer and other diseases. It can be painstaking work. Sometimes, says Andrew Feinberg, an epigeneticist at Johns Hopkins University in Baltimore, Maryland, the available techniques often pick up only "little biochemical shadows" of events going on at a particular location, while the complete set of players and their mechanisms remain mysterious. And even when you can identify an epigenetic molecule, says Tony Kouzarides, a molecular biologist at the University of Cambridge, UK, "you have to work out why it is there, and what it is doing there".

Nonetheless, epigeneticists have made remarkable progress over the past two decades. Their tool kit now includes advanced sequencing techniques, targeted antibodies and even laser cell sorting — and it should soon encompass ultrasensitive nanofluidic and nanopore sequencing methods. The community is also turning to advanced bioinformatics to cope with the sheer volume of data — especially the wealth of epigenetic information from the Encyclopedia of DNA ▶

Epigenetics
Made Simple

Products and Services for
DNA Methylation &
Hydroxymethylation Analysis

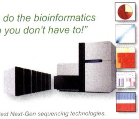
5-mC & 5-hmC Research Tools

- The most cited bisulfite technologies for DNA methylation analysis: easy-to-use single, 96-well & automated formats.
- Essential tools for your 5-hmC analyses.

Genome-wide
5-mC & 5-hmC Services

Simply send us your DNA from any species and receive the genome-wide methylation or hydroxymethylation epigenetic profile in a publication-ready format.

"We do the bioinformatics so you don't have to!"



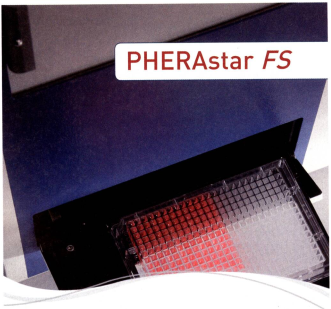
Powered by the latest Next-Gen sequencing technologies.

Zymo Research provides the most comprehensive offering of validated products for epigenetics research.

www.zymoresearch.com

ZYMO RESEARCH
The Beauty of Science is to Make Things Simple

THE
Epigenetics
COMPANY™

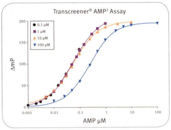


PHERAstar *FS*

The Gold Standard Microplate Reader
for High-Throughput Epigenetic Screening

Multi-mode HTS Microplate Reader that is perfect for studying epigenetic assays, including histone modification, DNA methylation and small interfering RNA. Unique Features:

- Two laser based excitation sources
  - UV Laser for TR-FRET/HTRF® measurements
  - Far Red Laser for AlphaScreen® Technology
- Tandem Technology – CCD spectrometer and filter based optics
- Simultaneous Dual Emission detection – Four PMTs
- Enhanced cell-based assays with Direct Optic Bottom Reading
- Onboard reagent injection for precise kinetic measurements
- All microplates up to 3456-well



Methyltransferases and Acetyltransferases can be measured using the Transcreener® AMP² Assay. Z' value of 0.91 is obtained with the PHERAstar FS.

3456 well plates

Meet us at **SLAS2013** in January in Orlando, Florida at **Booth #721**.
Find further information on www.bmglabtech.com

BMG LABTECH
The Microplate Reader Company

Alpha Technology includes AlphaScreen, AlphaLISA, and AlphaScreen SureFire. These assays are registered trademarks of PerkinElmer, Inc. HTRF is a registered trademark of CisBio International.

Elements (ENCODE) project, which this year released more than 1,600 genome-wide data sets covering more than 100 cell types[1].

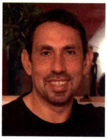
Technology development is now kicking into high gear as epigenetics researchers push to decipher the genome's many partners, and to deepen understanding of health and disease.

BEYOND A PRECIPITATING HEADACHE

The standard method used to study epigenetic histone modifications is called chromatin immunoprecipitation (ChIP), coupled with sequencing[2]. The basic idea is to shear DNA while it is still wrapped around the histones, use antibodies to capture specific protein–DNA complexes from the fragments, and then study which DNA sequences are attached to which proteins. The approach helps to unpick how the interactions are tuning genes — activating some, silencing others.

The technique has its drawbacks, however. Sriharsa Pradhan, an RNA biologist at New England BioLabs in Ipswich, Massachusetts, says that he is often unable to reproduce work from published epigenetics studies. "Most of the failures happen if the antibody is not good," Pradhan says. It might pick up too many DNA–protein complexes — "every Tom, Dick and Harry" in a sample — and so does not offer the resolution that scientists seek.

T. KOUZARIDES



"Whatever you see in one moment will change in the next."

Tony Kouzarides

Kouzarides agrees that the quality of the antibody matters greatly for ChIP and many other lab procedures. That's what led him to co-found Abcam, an antibody supplier with headquarters now in Cambridge, UK. The goal is exceptionally high quality, says Kouzarides, who is on Abcam's board of directors — but it is a constant struggle. "You are at the mercy of the rabbits," he says, referring to the animals used to generate the antibodies. "Some generate good antibodies, some generate bad antibodies" — and there is no predicting which is which.

Monoclonal antibodies could offer more reliability, says Kouzarides, because they avoid the problem of batch-to-batch variability. But for reasons that still aren't clear, he says, some of them do not work well for ChIP. For now, the field has to use animals to generate the antibody mixes useful for ChIP. "You have to put up with the unreliable nature of antibodies because it's the only way to do such experiments at the moment," he says.

Another drawback with standard ChIP is its bias, says Alan Tackett of the University of Arkansas for Medical Sciences in Little Rock. Although the technique lets scientists localize a specific protein acting on a genomic site, "you have to know what protein or histone modification you are targeting". And scientists need to have on hand an antibody that matches the protein of interest. So ChIP is not easily multiplexed to profile multiple areas of the genome at the same time.

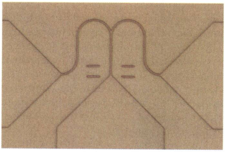
In response to this shortfall, Tackett and his Arkansas colleagues, along with scientists at the Johns Hopkins School of Medicine, have developed chromatin affinity purification with mass spectrometry (ChAP-MS)[3]. The approach involves cutting out a 1,000-base-pair region of a chromosome, purifying it and determining all the epigenetic changes that are present. The team has used the approach in yeast to detect different chromatin states, silenced genes and other regions in which genes are still active. And Tackett says that around ten other labs have begun exploring it, too.

He is now readying the technique for use in human cell lines and tissues. "We are working on the mammalian version and anticipate having that complete within the year," he says. One challenge for ChAP-MS is that the analysis requires 10^7 to 10^{10} cells, so Tackett and his colleagues are trying to lower that number. And Tackett is confident about the technology's promise. "We see this ultimately taking the place of ChIP in epigenetics labs," he says, with mass spectrometry being available through proteomics core facilities on campuses, he says.

Other scientists are proposing different alternatives to ChIP, which "is not a very efficient process", says Paul Soloway at Cornell University in Ithaca, New York. In addition to the challenges involved in sample processing, ChIP usually queries just one epigenetic mark at a time in a population of cells. That means that the results of multiple ChIP-seq experiments have to be aligned to determine if some cells have one mark and others have another, or if, perhaps, all cells have both.

Soloway wants to offer scientists greater resolution for ChIP analysis. He also wants the approach to be scalable, delivering detail and screening for multiple epigenetic marks in a single experiment. His answer is a nanofluidic device based on a silica wafer that is in the prototype stage and which comes in two formats[4]. One of them quantifies the molecules with at least one epigenetic mark. The other, a branched nanofluidic device, sorts and quantifies the molecules. Using fluorescent labels and optics-based sorting, the molecules are shunted to one chamber or another for later analysis, such as DNA sequencing. "Because silica is clear and non-fluorescent, we can make measurements of individual molecules using highly sensitive optics," he says.

P. SOLOWAY/CORNELL UNIV.



A nanofluidic device can sort through DNA molecules to find those with epigenetic marks.

Ultimately, Soloway would like to be able to go through whole genomes in a rapid, multiplexed way. He says that standard ChIP is still ahead of his technique because it can generate materials in the amounts needed for sequencing, whereas he still needs to get from single molecules to the pico- and nanograms needed.

Soloway believes that his technology will find a home in drug development, helping researchers to quickly and quantitatively characterize how drug candidates affect epigenetic marks. Clinical applications could include helping to monitor how patients fare when treated with epigenomic drugs, and identify how epigenetic marks vary during the course of a disease such as cancer, he says. In January, together with the Cornell engineers Harold Craighead and Stephen Levy who worked on the technology, he founded Odyssey Molecular in Ithaca, to commercialize the device.

FINDING OTHER MARKS

DNA methylation has important roles in cells, including the regulation of genes during development and disease. One of several methods used to find these sections of the genome is methylated-DNA immunoprecipitation, which uses an antibody that locates 5-methylcytosine, a methylated form of the DNA base cytosine.

A different approach targets methylated parts of the genome in 'CpG islands', which are characterized by a specific chemical bond between the DNA bases cytosine and guanine. In an analysis of methylation levels for 240,000 of the several million CpG islands in the ENCODE data, John Stamatoyannopoulos at the University of Washington in Seattle and his colleagues found a strong association between methylation and accessibility for genes to be read[5]. As Wendy Bickmore from the Medical Research Council Human Genetics Unit at the University of Edinburgh, UK, notes, the results support the idea that DNA methylation is blocked where the transcription factors that read DNA bind. This mechanism, she says, is relevant to the interpretation of disease-associated sites that show altered DNA methylation[6].

One widely used technique to determine DNA methylation patterns across a genome is bisulphite sequencing. The addition of bisulphite to DNA converts cytosine to uracil, but skips methylated cytosines, thereby allowing the methylation status of DNA segments to be determined through high-throughput sequencing. Many companies offer bisulphite conversion kits. "It's cheap enough now and there are statistical tools for understanding it, so there's no reason to use another method," says

Feinberg.

Yet detecting methylation is time-consuming, so scientists in academia and industry have been exploring ways to improve the approach. Some teams, including one at Osaka University in Japan and one at the University of Oxford, UK, are exploring the use of nanopores, tiny gates through which to run a DNA strand. And Pacific Biosciences, a sequencing firm in Menlo Park, California, is using tags to prepare single strands of DNA for high-throughput sequencing.

At Washington University in St Louis, meanwhile, Rob Mitra is leading an effort to be more precise in capturing methylation data, because this information can, for example, be an early sign of tumour development. Mitra and his team, including graduate student Maximiliaan Schillebeeckx, have developed a technique that uses lasers to separate out the cells of interest. He calls the technique laser capture microdissection–reduced representation bisulphite sequencing. Among the advantages, says Mitra, is that the technique covers "the majority of the CpG islands and it's relatively inexpensive".

Reduced representation bisulphite sequencing is similar to whole-genome bisulphite sequencing, but sequences only the parts of the genome that include CpG-dense regions. The technique uses enzymes to cut up purified genomic DNA into fragments that contain CpG islands. The fragments are then processed, and those of a certain size are subjected to bisulphite conversion, amplified and then sequenced.

The approach is geared to work on small amounts of DNA — perhaps even less than a nanogram — and in formalin-fixed, paraffin-embedded tissue, which is "typically not in as good shape as good fresh frozen DNA", Mitra says. This type of tissue fixation is typically used in biobank samples.

His technique could be a tool for researchers who work with specific cell types or with complex tissues, such as neurological samples, in which it is hard to isolate the cell type of interest, he says. The method also avoids the need for multiple labour-intensive purifications. And, he says, "at each point in space, you get a genome-wide profile of methylation, so now you can start to correlate methylation profiles spatially", Mitra says. A researcher can see, for example, if similar regions of complex tissue are methylated similarly. By coupling genome-wide methylation analysis with laser capture to isolate targeted cell populations, the tool can help researchers to address questions in these challenging tissues, he says.

EXPANDED REACH

Along with the flood of data that ENCODE brought to epigenetics came data standards, quality metrics, software tools and ways to convey how experiments are done, allowing comparisons between labs. This development has heightened awareness about the "good technologies" needed to study how the genetic code is put into action, says Adam Petterson, a senior scientist at Zymo Research in Irvine, California, which is one of many companies offering epigenetics services to academics as well as drug-discovery companies.

Such awareness is going to become ever more important as epigenetics grows to encompass not just multiple cell types, but multiple species. The modENCODE project (www.modencode.org) is mapping regulatory patterns in two frequently used model organisms, the fruitfly *Drosophila melanogaster* and the nematode worm *Caenorhabditis elegans*, and the Mouse ENCODE consortium is focusing on epigenomic mapping of the mouse. "A huge way to understand function is by comparative epigenomics," says Feinberg, who would like to see efforts across many more species.

These developments will inevitably require increased reliance on massive computation, says Kouzarides, who sees bioinformatics as a rate-limiting step in epigenetics. Researchers need ways to integrate and do global analyses of the emerging maps of epigenomic marks and their effects, as well as ways to do high-resolution analyses, preferably at the single-cell level (see page 27). Without such computational tools, Kouzarides says, "it's almost impossible to appreciate the complexity of the information".

For scientists who would rather not dig into the data themselves, Michael Snyder and his team at Stanford University in California have developed Regulome-DB (regulome.stanford.edu), an automated tool to explore non-coding regions of the human genome. Manolis Kellis at the Massachusetts Institute of Technology in Cambridge and his group have set up Haplo-Reg (www.broadinstitute.org/mammals/haploreg), a tool that helps to link non-coding variant patterns to possible clinical conditions.

B. BOSTON

"Now you can start to correlate methylation profiles spatially."
Rob Mitra

TRANSIENT DRUGS

The potential for clinical applications is an important motivator for epigenetics research. The transient nature of epigenetic changes gives drug developers and biomedical researchers reasons to dream about how their efforts might reverse changes that contribute to disease. "Those sorts of things that are more malleable are likely the things that we can target," Feinberg says.

Four drugs that act on epigenetic pathways have been approved by the US Food and Drug Administration (FDA), and the next wave of candidates is being readied in biotech and pharmaceutical companies. Kouzarides, for example, is looking at chromatin modifications and develops drug candidates that could right the wrongs in cancers in which, for example, epigenetic influences lead to the misregulation of cell pathways[7].

Targeting an aggressive form of leukaemia for which treatments are lacking, Kouzarides and his team have explored how to inhibit bromodomain and extraterminal (BET) proteins and remove them from chromatin. BET proteins belong to a class of epigenetic reader that targets histones, recruits multi-protein complexes to the spot where they attach and instructs cellular processes involved in reading genetic information.

The journey from the lab to the clinic is not usually quick, Kouzarides says. In this case, however, a candidate under development for inflammation was found to be applicable for the leukaemia. Now, the small-molecule inhibitor of the BET protein is in clinical development at GlaxoSmithKline, headquartered in London.

Kouzarides believes that chromatin-modification pathways are promising drug targets because they involve proteins interacting with other proteins. In the past, drugs have tended to target enzymes, and it has not been considered feasible to target protein–protein interactions with small molecules. But his work[8], along with that of others, has shown that it is possible to develop specific small molecules against the BET proteins that recognize a small epigenetic modification present on chromatin.

Constellation Pharmaceuticals in Cambridge, Massachusetts, is also exploring the BET family, as well as other enzymes that modify chromatin[9]. These therapies are going to be part of the second-generation epigenetic drugs that target specific modifications with a role in disease, explains Keith Dionne, the company's president and chief executive. The past, more coarse scientific understanding of chromatin has shifted to an appreciation of the "subtle distinctions" between chromatin states, explains James Audia, the company's chief scientific officer.

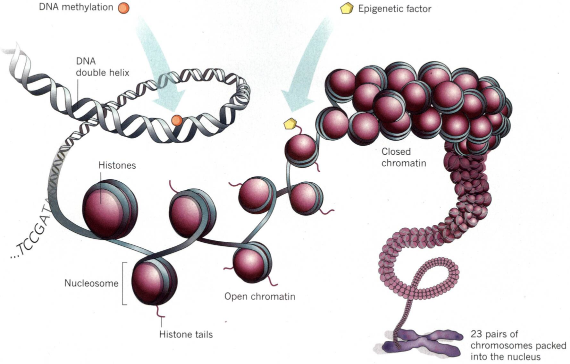
Earlier this year, Constellation and Genentech began collaborating on the development of inhibitors of BET proteins and another class of epigenetic modifier, the EZH2 chromatin-writers. These proteins seem to be part of a complex that represses gene expression; mutated versions have been linked to some cancers.

As Patrick Trojer, director of biology at Constellation Pharmaceuticals explains, cancers use chromatin modification to gain an advantage, for example to inactivate a pathway that creates room for unhindered tumour growth. As part of the company's drug-discovery programme, he and his colleagues develop techniques to study the details of chromatin changes. The understanding of chromatin biology is one of the company's strong suits, he says.

To support this application-based research, Trojer and his colleagues use a number of

POLYGAMOUS DNA

DNA works with many partners. DNA methylation, for example, influences the way that genes are expressed without changing the underlying DNA sequence, and other epigenetic factors bind to histones to control when chromatin complexes open up and allow their DNA to be read.



epigenetic techniques. ChIP-seq is a lab standard in which antibodies are "the key" to the technique, he says. But the company has also made histone mass spectrometry a priority, because it allows the scientists to query the chromatin changes without using antibodies and to query a number of modifications at once. The company set up an in-house high-throughput facility to screen for potential compounds.

Although other companies tend to outsource these tasks, the company wants to integrate findings about chromatin biology into drug discovery with an in-house suite of tools that includes mass spectrometry and biophysics analyses, Dionne explains.

Another Cambridge-based epigenetics company, Epizyme, focuses on a family of proteins called histone methyltransferases. These epigenetic modifiers act on histones, by catalysing the transfer of methyl groups onto specific positions in the protein. The company has partnerships with the pharmaceutical companies GlaxoSmithKline; Celgene Corporation in Summit, New Jersey; and Eisai in Woodcliff Lake, New Jersey, as well as the Leukemia and Lymphoma Society in White Plains, New York, and the Multiple Myeloma Research Foundation in Norwalk, Connecticut. So far, 96 histone methyltransferases have been identified in humans, says Robert Copeland, Epizyme's chief scientific officer. "We believe there are at least 20 of those enzymes that are high-value targets for human cancers."

The company's goal is to find a molecule that blocks an enzyme active in an epigenetic pathway but not its nearest neighbours, he says. It is a selectivity that has been difficult to come by in the development of biotherapeutic drugs.

Copeland believes that epigenetics drugs fit into a trend of defining a cancer not by its anatomical location but by its molecular profile, which includes epigenetic signatures. Like many companies in this field, he and his colleagues mine the publicly available databases, noting that many genetic alterations in epigenetic pathways are found in human cancers.

Kouzarides believes that many cancer cells will be very vulnerable to epigenetic drugs because they rely on only one or two epigenetic pathways, whereas normal cells draw on several pathways for their functions. At the same time, he believes that epigenetics researchers and technology developers will still want to develop and refine experimental methods, for example to explore the three-dimensional structure of epigenomic events, to see how chromatin is changing throughout the genome. "It's very difficult to look at chromatin itself," he says. "Technology still has to evolve to look at *in vivo* chromatin effects." The available epigenetic data are "extensive, but still a very small snapshot" of epigenetic changes, he says. They represent a situation at a specific time in a specific cell.

Epigenetics might find its way into preventive medicine, too. Scanning the epigenome could be a way to detect disease well before symptoms arise. The blood pricked from the heels of newborn babies is one way to begin. In many countries, the blood spots are placed on Guthrie cards and stored indefinitely by hospitals and health-care systems. Scientists at Queen Mary, University of London are exploring how DNA methylation patterns change between newborns and in cells from the same children when they are three years old. Differences in epigenetic marks could be clues to health.

If the sequencing companies are betting right, then genome sequencing could become commonplace for many patients, perhaps even part of an annual physical examination. An epigenetic read-out, updated at regular intervals, might be an important companion file to that genome sequence.

But this type of progress depends on deeper understanding of epigenetic mechanisms and technology that has yet to evolve, Kouzarides says. Because epigenetic events change constantly in the cell, "whatever you see in one moment will change in the next", he says. ■

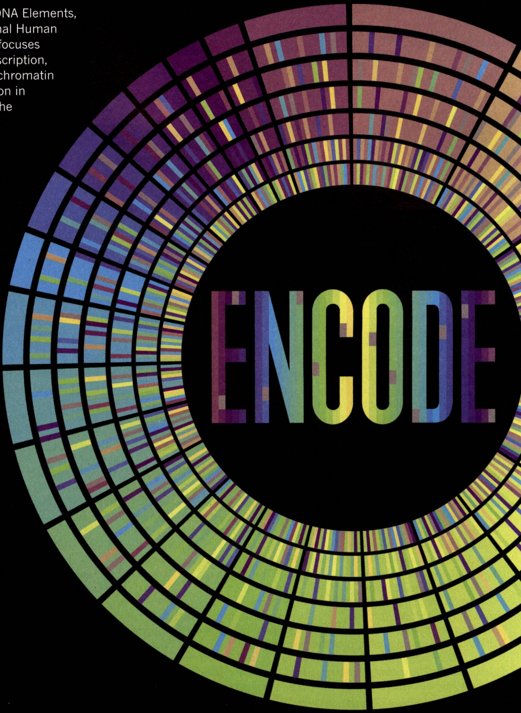
Vivien Marx *is technology editor at* Nature *and* Nature Methods.

1. The ENCODE Project Consortium *Nature* **489,** 57—74 (2012).
2. Park, P. J. *Nature Rev. Genet.* **10,** 669–680 (2009).
3. Byrum, S. D., Raman, A., Taverna, S. D. & Tackett, A. J. *Cell Rep.* **2,** 198–205 (2012).
4. Cipriany, B. R. *et al. Proc. Natl Acad. Sci. USA* **109,** 8477–8482 (2012).
5. Thurman, R. E. *et al. Nature* **489,** 75–82 (2012).
6. Ecker, J. R. *et al. Nature* **489,** 52–55 (2012).
7. Dawson, M. A. & Kouzarides, T. *Cell* **150,** 12–27 (2012).
8. Dawson, M. A. *et al. Nature* **478,** 529–533 (2011).
9. Mertz, J. A. *et al. Proc. Natl Acad. Sci. USA* **108,** 16669–16674 (2011).

nature | ENCODE

ENCODE, the Encyclopaedia of DNA Elements, is a project funded by the National Human Genome Research Institute and focuses on identifying all regions of transcription, transcription factor association, chromatin structure and histone modification in the human genome sequence. The main findings of ENCODE are now presented in 30 papers published simultaneously in *Nature*, *Genome Research* and *Genome Biology*.

Access videos, Features and the collected research papers, and explore the thematic threads that run through them via the *Nature* ENCODE explorer or the *Nature* ENCODE app.



Produced with support from

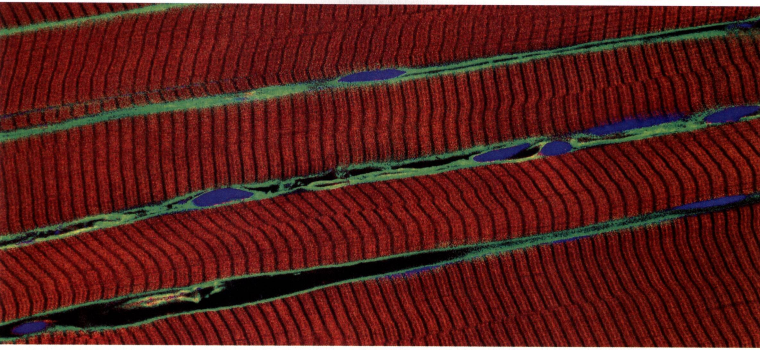
illumina®

Available on the App Store

nature.com/ENCODE

nature publishing group npg

nature
medicine



ePoster on Metabolic Syndrome

Diabetes and obesity are often referred to as the twin epidemics, owing to their high prevalence and mechanistic relatedness. A key concept that has emerged over the past few years is that different organs have different, but integrated, roles in the pathogenesis of the metabolic syndrome. This ePoster, a collection of figures prepared by leaders in the field, presents a comprehensive view of the molecular pathophysiology of the metabolic syndrome at the systems and cellular levels. Published originally in 2011, the 2012 version features a new organ -- the pancreas -- plus updates for the existing organs to reflect key advances of the past year.

Access the ePoster FREE online!
www.nature.com/nm/poster

Image credit: Tom Deerinck, UCSD

Sponsored by

Takeda

nature publishing group npg

Principles of Biology
by nature EDUCATION
Principles of Biology
Principles of Biology
Principles of Biology
Bringing biology to life

To learn more about *Principles of Biology*, the first interactive textbook from Nature Education, visit **www.nature.com/principles**

nature publishing group npg

CAREERS

WOMEN IN SCIENCE Campaign to boost women in science seeks videos **p.151**

FACULTY 'Credit' swap programme aims to provide greater flexibility **p.151**

NATUREJOBS For the latest career listings and advice **www.naturejobs.com**

M. WISNIEWSKA/SHUTTERSTOCK

FOOD SCIENCE

A smorgasbord of opportunity

Food science is fun — and boasts better job prospects than many other industry niches.

BY ALLA KATSNELSON

Tara McHugh recalls the JustFruit bar with more than a dollop of pride. In her early days at the US Department of Agriculture (USDA), McHugh led development of the technology for making 100% fruit bars, and then helped apple and pear growers to launch a company, Gorge Delights, to manufacture them. "We wanted to get people to eat more fruit," she says. The firm, based in North Bonneville, Washington, still sells the product a decade later, and even now it occasionally checks in with manufacturing questions, says McHugh, leader of processed-foods research at the USDA's site in Albany, California. Her team has also taken on projects such as helping Newgem Foods in Stockton, California, to develop sandwich wraps made from vegetables, and working with Monterey Mushrooms in Watsonville, California, on a growing process that ups the vitamin D content of fungi.

McHugh leads about a dozen scientists who study how food processing affects nutrition, and devise food processes that boost health. "One of the things I've liked most about my job is the ability to partner and take the basic research we develop into commercialization," she says.

The science of food is a great career choice for scientists of a practical bent, says Charles Shoemaker, academic adviser for undergraduate and graduate students in the food-science department at the University of California, Davis. "We're all familiar with eating, all our lives," he says. Food science "is a way to go into a scientific career and stay anchored to something you're really comfortable with". As a profession, it has long been something of a secret, says Shoemaker, but that is starting to change — owing in no small part to the rise of celebrity chefs and television programmes about food and cooking. "It sounds kind of silly in a way, but [US television channel] the Food Network has added a really serious degree of legitimacy to food science," he says.

The field encompasses a wide range of scientific disciplines: the chemistry of how ingredients interact; the microbiology of harmful and helpful bacteria; the biology of the sensory processes that underlie smell and taste; nutritional assessment; and the engineering involved in creating measurement tools or manufacturing, to name a few. Researchers might work with processes that make a type of food tastier, healthier or cheaper; oversee quality standards or compliance with government regulations; or specialize in studying an elemental component of food, such as proteins or carbohydrates.

MIX IT UP

More than in other disciplines, says Shoemaker, food scientists must be able to collaborate, because so many steps, components and ideas go into developing a food product. "You'll take on a project, and at some point you'll have to deal with food-safety issues, chemistry to deal with the stability of a product, engineering in scaling up production," he says. "You're not necessarily going to be an expert in those areas, but you need to be able to converse with relevant people and understand their issues and how they reflect on your product."

Most people enter the job market after completing dedicated undergraduate or graduate degree programmes in food science. The curricula of such programmes often underscore the interdisciplinary approach: Wageningen University in the Netherlands, for example, runs a respected master's programme with specializations in areas such as food biotechnology, sustainable food processing and ingredient functionality. The course features on a list of 52 graduate programmes that meet the educational standards of the Institute of Food Technologists (IFT) in Chicago, Illinois, a professional group allied with food-science organizations around the world.

Most academic programmes require students to complete internships to get more substantive research experience in work environments. A placement might be at a multinational giant such as PepsiCo, based in Purchase, New York, Heinz; based in Pittsburgh, Pennsylvania; or Cadbury, based in London. Or it might be ▶

Tara McHugh helped to develop vegetable-based sandwich wraps.

P. GREB

▶ in a food biotechnology start-up with few staff members; a government facility such as the US Food and Drug Administration or the European Food Safety Authority; or an academic lab.

Even in academia, much food research is funded by companies or trade groups. On the plus side, that gives young researchers plenty of options for funding, and it increases fluidity between sectors, so scientists working in industry may have the opportunity to return to academia.

TOO MANY COOKS

The heavy industry involvement can also bring challenges, says Megan Clements, who is finishing a food-science PhD at the University of California, Davis, and has already taken on her first job, at Hampton Creek Foods, a start-up in San Francisco. Her dissertation looks at how to measure almond quality and is funded by the Almond Board of California, based in Modesto. "It was a marriage of convenience in some ways, but it was a good marriage nonetheless," she says. However, she cautions that the goals of some industry-proposed projects are too narrow to give students the research experience they need to satisfy their academic requirements and add value to their CVs. Industry agendas can also be fickle. "If suddenly they are not interested in what you originally started with, they might ask you to switch your topic two years in," she says, adding that graduate students should recognize that academic advisers can also have vested interests linked to industry partners that might conflict with the student's needs.

The long arm of industry brings solid job prospects — a particular perk in the struggling economy. Unlike the pharmaceutical industry, which has haemorrhaged research staff over the past few years, food-science recruiters and long-time researchers say that the sector has been touched only lightly by the economic crisis. According to the IFT, earnings for new recruits had dropped by 2009, but have been creeping back up. The organization's annual salary survey showed that in 2011, US food scientists earned a median of $87,000 (M. E. Kuhn *Food Technol.* **66,** 26–38; 2012). Entry-level salaries start at about $40,000; for jobs such as vice-president of global research and development at a multinational company, remuneration can easily reach $250,000, says Susan Canarick, senior executive recruiter at Opus International, a recruitment firm based in Deerfield Beach, Florida, that specializes in food-science positions.

> *"Nearly everyone who graduates has a job, and many are choosing from two or three."*

In Europe, master's-degree holders can expect entry-level salaries of €30,000 (US$39,000) or more for positions in product development or food safety at medium-to-large companies. Researchers with a PhD in an area such as dairy or bakery science could get €45,000–50,000 for an entry-level product-development post, and double that with 5–10 years of experience, says Jan-Willem Broekhoven, a managing partner at recruiting firm SIRE Life Sciences, based in Amsterdam. Jobs with more of a commercial bent — in which researchers must provide technical assistance to customers or keep abreast of regulatory issues, and must have higher-level expertise in a niche topic such as starches or cocoa — might pay €65,000–80,000, he says. "Nearly everyone who graduates has a job, and many are choosing from two or three," says Ralf Hartemink, director of the food-science programme at Wageningen University.

INGREDIENTS FOR SUCCESS

Product development is a common area for researchers to break into industry. Research and development at multinational companies with developed product lines may be tame, says Shoemaker, but there is plenty of innovation and entrepreneurship in the start-up world, which is particularly well-established in California. "It's what the biotech world was like ten years ago," he adds.

For instance, Hampton Creek hired Clements to develop a plant-based product with the right emulsion properties to replace egg in products such as mayonnaise. The company believes that taking eggs out of food preparation improves safety by lessening exposure to pathogens such as *Salmonella* and bird flu, as well as significantly reducing costs and increasing sustainability.

Other areas with good job prospects are nutrition and food safety. Consumers are increasingly demanding healthy foods that contain reduced amounts of fats, calories or sodium, or carry benefits such as probiotics to improve digestive health or fibre to reduce the risk of heart disease. The health claims of such products must be tested, and some items must be reformulated for specific ends.

A growing awareness of food safety has also created a need for researchers to monitor microbial and chemical contamination, and design safer food-development processes. US industry and government agencies are likely to add jobs to ensure compliance with the Food Safety Modernization Act, passed in 2011.

Such highly structured regulatory work is one of two types of government food-science jobs, notes Shoemaker. The other is research geared towards health and the public good. McHugh considers her post at the USDA to be a cross between an academic and industrial position. "We still publish manuscripts and are evaluated largely on our fundamental research, but we also have a unique ability — and it's part of our remit at the USDA — to transmit these technologies to commercialization," she says.

Although specialized academic programmes provide the easiest entry point into a food-science career, there are opportunities for researchers in almost any basic biological, agricultural or engineering field who have a love of food. Jennifer Kimmel, a protein chemist at Kraft Foods Group in Glenview, Illinois, never expected to stumble into a food industry job. But after spending five years as a postdoc studying enzyme kinetics at the University of Missouri in Columbia, she grew worried about getting funding. "I had to get a job really, really fast," she says. With her postdoc drawing to a close, she saw an advertisement for a position at Kraft that asked for experience with enzyme kinetics. She decided to apply, and got the job.

Kimmel's grounding in basic research has turned out to be a strong asset. It allows her to help product developers and other scientists to predict how proteins would work in particular products. That means understanding "not only the phenomena that are supposed to happen, but also the ones that aren't supposed to happen", she says. "No one wants a cheese that doesn't melt." ■

Alla Katsnelson *is a freelance writer based in New York.*

COLUMN

Career fast-forward

An early stab at leadership can help a young scientist to imagine heading a group, argues **Katharina Genreith**.

I am a 26-year-old PhD student. I am also an independent group leader — of a sort. I acquire my own funding, and run a lab with eight team members.

How could such a young student have such a serious responsibility? I am taking part in the Life-Science Lab: an unconventional education initiative at the German Cancer Research Center in Heidelberg. The programme gives students from high-school to PhD level the opportunity to join independent, mixed research groups and work on scientific projects of their choosing — in their own labs.

Students get the chance, in essence, to fast-forward their careers and see what it is like to be two steps ahead on the academic ladder. I have been given a glimpse of my future, working as a responsible group leader on synthetic biology.

Apart from our remarkably low mean age, my team and I are just like any bunch of scientists. What we might lack in experience and maturity, we have tried to offset with enthusiasm, hard work and the occasional chocolate cake at lab meetings. Our project uses genetically modified *Escherichia coli* bacteria to make living ultraviolet-light sensors integrated into jewellery pieces, to alert people when the sun has started to cause damage to their DNA.

COMPETITIVE DRIVE

Despite our enthusiasm, our productivity had been flagging. It can be difficult to convince volunteer team members to sacrifice their Sundays in the absence of deadlines or external incentives. But productivity got a boost when, with fellow Life-Science Lab participant Dominik Niopek, we joined the International Genetically Engineered Machine competition (iGEM). Suddenly, we had access to thousands of DNA parts (virtual and actual) from iGEM's Registry of Standard Biological Parts, and had the outstanding incentive of measuring ourselves against teams from all over the world. At the same time, however, we were exposed to clear project requirements, deadlines and the small problem of acquiring US$10,000 in funding for lab disposables and expenses for travelling to the competition finals in Greenfield, Indiana. I felt the pressure of being in charge.

The work demanded personal sacrifices, caused sleep deprivation and meant that we had to cope with high expectations from the sponsors and ourselves. But it was worth it. Ten weeks after our team signed up for the iGEM competition, we had raised about US$13,000 in funding from biotechnology companies, government education funds and generous grandmothers.

The experimental results were better than we had dared to hope. At the iGEM high-school competition jamboree, the judges awarded us not only the main prize — a green BioBrick trophy in the shape of a Lego brick — but also five out of ten special prizes.

JUMPING INTO THE DEEP END

Changing my perspective — from PhD student to group leader — gave me the opportunity to develop a scientific project from beginning to end, with responsibility for both the research and my team. I gained practical experience of managing a project, reconciling competencies and team dynamics, assigning tasks and providing constructive feedback. I also saw my shortcomings, including my tendency to grow impatient, which raised anxiety levels.

But most importantly, the career fast-forward taught me to come up with my own project ideas rather than relying on the expertise of my supervisor. I discovered that what I really love about science is the creative process and the possibility of making an impact on society. The experience reinforced my desire to pursue a research career.

By becoming group leaders, early-career scientists have a chance to broaden their horizons and acquire skills such as project organization, mentoring and coordination of a research team. And student-led research teams would work well in nearly any branch of science. The only things really required are a genial, enthusiastic group of students, lab space, the courage to stray off the beaten path — and the occasional chocolate cake. ■

Katharina Genreith *is a PhD student in cancer biology at the German Cancer Research Center in Heidelberg, Germany. She worked on the Life-Science Lab initiative.*

WOMEN IN SCIENCE

Video competition

In June, the European Commission released a short video intended to attract women to science careers as part of its Science: It's a Girl Thing campaign (http://science-girl-thing.eu). The widely maligned, sexually suggestive clip — created by an external production company— depicted flirtatious models strutting through a lab setting, ostensibly to help to make cosmetics. Criticized for trivializing the issue, it was soon taken down. Now the European Science Foundation in Strasbourg, France, has launched an online contest asking for members of the public from around the world, including scientists, to submit a video that the European Commission might use in its continuing women-in-science campaign. Entries must be submitted by 19 November and be no more than one minute long; they can be uploaded on either YouTube or Facebook. One prize will go to the video that earns the most votes from viewers between 19 and 28 November, and a jury will select two further winners. Each winning team will receive €1,500 (US$1,950).

FACULTY

Boosting flexibility

In an effort to retain faculty members by improving their work–life balance, Stanford University School of Medicine in California has created a credit-banking system. Faculty members will receive credits for the time they spend on mentoring, committee work and similar tasks; they can cash in the credits for gift certificates at businesses approved by the medical school to help them with tasks such as editing manuscripts, writing grant applications, housework or meal preparation. The system aims to put a formal value on time spent on non-research activities. Faculty members "can buy back time that they would have spent at home to help them elevate their role in the lab", says Jennifer Raymond, a neurobiologist at Stanford and associate dean for faculty development. "We're hoping to demonstrate that we can increase faculty flexibility and sanity," she adds. "We also believe that we'll increase retention and recruitment." The pilot programme will run for a year, after which administrators will compare faculty productivity and workplace satisfaction before and after the launch. If the pilot programme is deemed to be successful, it will be expanded to faculty members across the university.

MAKE YOUR RESEARCH GLOBAL!

10th Anniversary 2002 2012 CRG Centre for Genomic Regulation

ncbs national centre for biological sciences

inStem

NCBS - inStem - CRG Post-doctoral Fellowships

Joint Post-doctoral Fellowship Program between National Centre for Biological Sciences (NCBS), the Institute for Stem Cell Biology and Regenerative Medicine (inStem), Bangalore, India, and the Centre for Genomic Regulation (CRG), Barcelona, Spain

Deadline for applications:
15th November, 2012

Online application:
www.crg.eu/ncbs-crg_call

For further information:
sarina@ncbs.res.in
imma.falero@crg.eu

Positions open in these labs:
James Sharpe - Sandeep Krishna - Matthieu Louis - Thomas Graf Mukund Thattai - Toni Gabaldón - Jérome Solon - Jyotsna Dhawan Mara Dierssen - Ben Lehner - Raghu Padinjat - Pia Cosma Guillaume Filion - Isabelle Vernos - Madhu Venkadesan - Pedro Carvalho

Members of the CRG Board of Trustees:

Generalitat de Catalunya

Universitat Pompeu Fabra

NCBS is part of:

tifr

Centre for Genomic Regulation
(CRG), PRBB Building, Dr. Aiguader 88, 08003 Barcelona (Spain) – Ph: +34 93 316 01 00 – www.crg.eu

National Centre for Biological Sciences
(NCBS), Tata Institute of Fundamental Research, GKVK, Bellary Road, Bangalore 560065 (India) -
Ph: +91 80 23636460 - www.ncbs.res.in

W230304R

---

imb Institute of Molecular Biology

is recruiting

Research Group Leaders (#CNRGL04)

The **Institute of Molecular Biology (IMB)** Mainz, Germany, is a new research center of excellence situated on the campus of the University of Mainz and integrated into its scientific community. IMB is generously funded by the Boehringer Ingelheim Foundation with 100 M Euro. Projects at the institute are interdisciplinary and focus on epigenetics, DNA repair, development and the interfaces between these fields. Scientists at IMB are supported by strong Core Facilities.

We seek Group Leaders investigating nuclear function and gene regulation from the molecular to the organismic and systems biology levels, especially in the areas of:

- **RNA Biology**
- **DNA Repair**
- **Chromatin Biology**

We seek outstanding, collaborative candidates with an innovative research profile in the aforementioned fields and a strong publication record. Positions are initially for 5 years and can be renewed. Recruitment packages include an attractive salary, generous start-up funds and a 5-year budget for personnel and consumables. There are no teaching obligations, but teaching opportunities are available at the University, e.g. for individuals wishing to obtain a German Habilitation.

To apply, please send your cover letter, CV and a concise description of your research interests and future research plans to personnel@imb-mainz.de, quoting the reference #CNRGL04.

Please also arrange for 2 letters of recommendation to be sent to personnel@imb-mainz.de. For further information on IMB, please visit www.imb-mainz.de. Questions can be addressed to IMB s Director, Christof Niehrs (c.niehrs@imb-mainz.de). The **deadline for applications** is **20 January 2013. Interviews** will take place in Mainz on **25-26 February 2013.**

W230288R

MAX PLANCK INSTITUTE FOR **BIOLOGY OF AGEING**

Since its founding in 2008, the Max Planck Institute for Biology of Ageing in Cologne, Germany, has become a hub within a network of research institutes constituting a unique cluster in the field of ageing research worldwide.

The institute will move into a new building with state-of-the-art laboratories and facilities, located on the Campus of the University of Cologne and the Cologne Medical Center.

Fundamental questions on the nature of ageing and longevity are being investigated in various model organisms: Why do organisms age? What biological processes determine lifespan? And what are the roles of genes and environment? Researchers at the institute currently represent almost 30 different countries. The MPI is expected to grow to a size of 350 employees. There will be four directors with corresponding departments and four Max Planck Research Groups ready to welcome junior scientists of exceptional talent.

IMPRS-Age Graduate School
Investigating the nature of ageing and longevity

In close collaboration with the graduate school of the CECAD Cluster of Excellence, also based in Cologne, the International Max Planck Research School IMPRS-Age provides a first-class education in the field of ageing and age-related diseases. Outstanding PhD students will receive training in biological and medical disciplines, gain experience in the use of diverse model organisms as well as state-of-the-art technologies, and develop communication skills essential for a successful scientific career.

For a list of current vacancies, please visit **www.age.mpg.de**

W229074R

NDM SCIENTIFIC LEADERSHIP FELLOWSHIPS
at the Wellcome Trust Centre for Human Genetics

UNIVERSITY OF OXFORD

The Wellcome Trust Centre for Human Genetics, Roosevelt Drive,
Headington, Oxford, OX3 7BN

Salary scale: Grade 9: £42,883 - £49,689 with a discretionary range to £54,283 p.a.

Applications will also be considered for appointment at the appropriate salary range of Grade 8: £37,012 to £44,166 per annum, with a suitable adjustment to duties and responsibilities.

Applications are invited for NDM Scientific Leadership Fellowships at the WTCHG.

The fellowships are offered for five years and are intended to suit outstanding early career scientists with typically between four and eight years postdoctoral experience who would, once established, be in a strong position to enter one of the UK's established fellowship schemes funded by external bodies, such as the Wellcome Trust, the MRC, the Royal Society, the BBSRC, CRUK, or the BHF.

The WTCHG, an interdisciplinary centre of 470 scientists, is a world-leading research institute for the study of the genetics of common human diseases, structural biology and statistical genetics. We are looking to expand our range of expertise and would be particularly attracted to recruiting new group leaders interested in an aspect of:

- Epigenetics, gene regulation, non-coding RNA, or meiotic recombination
- The genetics of a particular human disease

It is anticipated that one appointment will be made in each of these areas.

A generous start-up package will fund personal salary, research staff and consumables. There is flexibility over starting dates, but successful applicants are anticipated to take up post by early autumn 2013.

In addition to these two posts, we are also interested in expressions of interest from established senior researchers wishing to relocate to the WTCHG. Funding to assist is negotiable, and substantial supplements are potentially available for those relocating from overseas. Informal enquiries from senior researchers should be addressed to the Business Manager, Peter Bond, by email: peter.bond@well.ox.ac.uk

The positions are fixed-term, for five years in the first instance and funded by the Wellcome Trust and the Nuffield Department of Medicine.

To apply and for further details, including a job description and person specification, please click on the link below:

https://www.recruit.ox.ac.uk/pls/hrisliverecruit/erq_jobspec_version_4.jobspec?p_id=105163

Only applications received before **12:00 midday** on **Monday 19th November 2012** can be considered. Interviews will be held on 10th & 13th December.

Please quote reference **105163** on all correspondence. You will be required to upload a CV and supporting statement as part of your online application.

U230037R

TGAC
The Genome Analysis Centre™

BBSRC EEDA East of England Development Agency Greater Norwich Development Partnership

Building Excellence in Genomics and Computational Bioscience

Chair, Scientific and Legal Non-Executive Directors and Trustees

The Genome Analysis Centre (TGAC) is a vibrant research institute focused on the application of state of the art genomics and bioinformatics to advance plant, animal and microbial research and promote a sustainable bioeconomy. It is located with five other partners on the Norwich Research Park, a new bioscience cluster with an excellent reputation in plant and microbial sciences, interdisciplinary environmental science and food, diet and health.

Chair - The Trustees of TGAC are seeking a high profile internationally renowned scientist to lead them through a demanding period of transition to follow a new scientific vision.

Scientific Trustees (2) - You will have demonstrable relevant scientific skills in the genomic, bioinformatics or systems biology fields and previous experience of working as a director of a high profile board or committee.

Legal Trustee – You will be operating at a senior level in a commercial legal area of practice, and have an interest in TGAC's science.

As the Chair or Non-Executive Director and Trustee of the Charity you will use strategic leadership skills together with integrity to lead TGAC. You will be prepared to devote at least 10 days a year to the role.

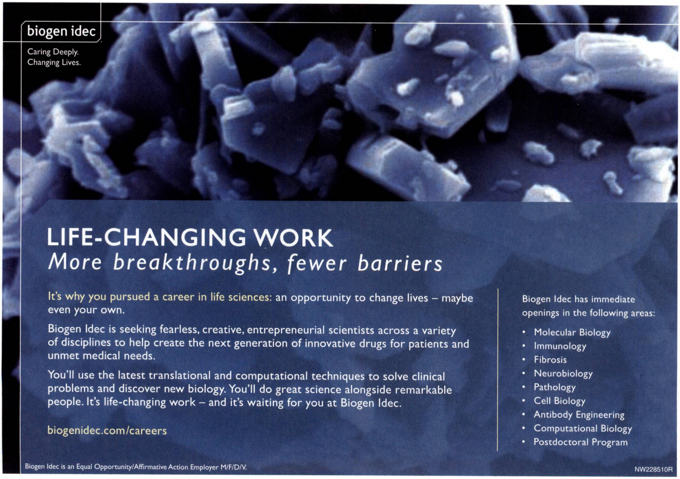
If you are committed to contributing at Trustee level to the governance of a cutting edge scientific institution and would like to apply please download details of this position from www.cfappointments.com and upload your CV and a letter of application.

Closing date: 11 November 2012
Initial interviews in London: w/c 12 - 23 November 2012
Final Interviews in Norwich: 12, 13 and 14 December 2012

For more information about TGAC visit **www.tgac.ac.uk**

cfA A PROSPECTS COMPANY
Executive Search and Selection | Recruitment

U230058R



biogen idec
Caring Deeply.
Changing Lives.

LIFE-CHANGING WORK
More breakthroughs, fewer barriers

It's why you pursued a career in life sciences: an opportunity to change lives – maybe even your own.

Biogen Idec is seeking fearless, creative, entrepreneurial scientists across a variety of disciplines to help create the next generation of innovative drugs for patients and unmet medical needs.

You'll use the latest translational and computational techniques to solve clinical problems and discover new biology. You'll do great science alongside remarkable people. It's life-changing work – and it's waiting for you at Biogen Idec.

biogenidec.com/careers

Biogen Idec has immediate openings in the following areas:

- Molecular Biology
- Immunology
- Fibrosis
- Neurobiology
- Pathology
- Cell Biology
- Antibody Engineering
- Computational Biology
- Postdoctoral Program

Biogen Idec is an Equal Opportunity/Affirmative Action Employer M/F/D/V.

NW228510R

WESTFÄLISCHE
WILHELMS-UNIVERSITÄT
MÜNSTER

Zertifikat seit 2008 audit familiengerechte hochschule

The Faculty of Biology at the University of Münster invites applications for a tenured position at the Institute for Neuro- and Behavioral Biology starting March 2014 as a

W2-Professor in Molecular Behavioral Genetics

The successful candidate should be able to represent the areas of neurobiology and molecular behavioral genetics and must demonstrate internationally recognized research accomplishments and a commitment to innovative teaching. Candidates should complement existing lines of research at the Institute for Neuro- and Behavioral Neurobiology. Applicants with a strong background in the areas of behavioral biology of model organisms, cellular principles of neuronal function, molecular neurogenetics or optogenetics are particularly welcome. The positon is expected to contribute to teaching in the BSc, MSc and PhD programs including lectures and practical courses. Interactions with existing collaborative research programs (SFB 629, SFB TRR 58, Excellence cluster „Cells in Motion" and the Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience) and participation in the academic self-administration is expected.

Prerequisite for the application are scientific achievements as a junior professor, from a postdoctoral lecture qualification (habilitation), as a research scientist at a school of higher education/university, non-university institute, industry, administration, or other fields of society within or outside Germany.

The Westfälische Wilhelms-University Münster is an equal opportunity employer and is committed to raising the proportion of women scientists in senior positions. Consequently, we actively encourage applications from suitably-qualified women. Women with the qualifications and disciplinary expertise required will be preferential considered. We also welcome applications from candidates with severe disabilities who, with suitable qualifications, will be preferential considered.

Applications must be submitted electronically as a PDF file that includes a CV, documentation of scientific career, teaching experience, list of publications, third-party funds acquired and copies of the five most important publications.

Documents should be submitted by **December 15, 2012** to:
Prof. Dr. Dirk Prüfer; Dekan des Fachbereichs Biologie; Westfälische Wilhelms-Universität Münster; Schlossplatz 4; 48149 Münster

www.uni-muenster.de

W230281R

Director at Centre for Research in Agricultural Genomics Consorcio

Barcelona

crag CENTRE DE RECERCA EN AGRIGENÒMICA

The Centre for Research in Agricultural Genomics (CRAG) is a CERCA Research Institution supported by a consortium of IRTA, Consejo Superior de Investigaciones Científicas (CSIC), Universitat Autònoma de Barcelona (UAB) and Universitat de Barcelona (UB). CRAG is a centre devoted to cutting-edge research in the molecular basis of genetic characters of interest in plants and farm animals and in the application of molecular approaches for the breeding of species important for agriculture and food production. CERCA Institution is the public Catalan network of research Institutes.

Candidate profile and contact information

CRAG is seeking applicants with a distinguished record of scientific excellence and the innovative thinking necessary to lead a dynamic organisation. He/She should have managed international projects or consortia, in particular European Framework Programme R&D projects. Salary will be commensurate with experience. Candidates must have in-depth knowledge about the Spanish Scientific environment.

Applicants should send a CV and a cover letter by e-mail to the Director of CERCA Institution at **applications@cerca.cat.** The appointment will be for a period of 4 years, which could be extended up to a maximum of 8 years.

Deadline, 11th November 2012

More information at **http://cerca.cat/en/crag-open-call/**

Institució CERCA Centres de Recerca de Catalunya CSIC IRTA UAB UB

W230275R

Bioinformatics Positions in Neuroscience & Personalized Medicine

to Study (Epi)Genomic Bases of Neural Functions, Epilepsy, Neurological Disorders

INTELLIMEDIX

Advanced PhD-level positions are available immediately for computer scientists and computational biologists to study Genomic Bases of Neural Functions. Work will be multifaceted and include assembly and annotation of RNA-seq and ChiP-seq data from diversity of neural circuits and even from single-cells, as well as identification of novel gene regulatory pathways and mechanisms underlying neural circuit organization, epilepsy (e.g., Dravet, GEFS+, Lennox Gastaut) and neurodegeneration. We will also perform genome sequencing of a subset of the population to identify additional "targets" and develop novel diagnostic and treatment approaches. Previous experience in genome assemblies, annotation and analysis of next generation sequencing (NGS) data (e.g. genome-/exome- sequencing and whole genome association studies (WGAS), DNA methylation/methylome, chromatin, or system biology) are preferable.

These are ideal positions for motivated, self-driven, and adventurous PhD-level scientists who would like to be integrated into interdisciplinary programs, and interact with both fundamental and clinical researchers from a location that provides excellent access to Florida's oceans. The job includes unprecedented opportunities for traveling and collaboration across the globe.

Preferable Qualifications: Ideal candidates should have 1) a Ph.D./MS in Bioinformatics/Genome or Computer Science or related field. 2) Experience in genome and/or transcriptome assembly software and/or mathematical modeling. 3) Strong skills in at least one programming language (Python, JAVA or C/C++, MATLAB, R, PERL) and/or large-scale analysis and databases.

Contact: llm@intellimedix.com. Please provide your CV, names of three references and evidence of your experience in the field of bioinformatics and/or computational biology and/or computer sciences. A proficiency in database development and maintenance, together with an ability to write custom scripts and/or algorithm development for informatics/visualization is a plus. Additional information can be found at http://intellimedix.com/index.html

NW230211R

Bringing jobseekers and recruiters together.

Academic • Industry • Government

Search now at **naturejobs.com**

naturejobs
making science work

Department of Medical & Molecular Genetics, Division of Genetics & Molecular Medicine, School of Medicine

KING'S College LONDON

King's College London

This is a call for expressions of interest for a leadership role in the Department of Medical and Molecular Genetics. In addition to a Senior Professorial appointment, there are opportunities for appointments at Senior Lecturer and Lecturer levels in this strong and growing Department. The Division of Genetics and Molecular Medicine also includes the St John's Institute of Dermatology, the Department of Twin Research and Genetic Epidemiology and the King's Centre for Stem Cells and Regenerative Medicine. The Division has strong research programmes in both rare and complex genetic disease as well as cancer genetics. It has formal links with the Biomedical Research Centre and houses the BRC Genomics Laboratory.

The Department wishes to attract researchers with a proven track record in state of the art genomic technologies who wish to establish a successful research programme in functional genomics that is likely to incorporate one or more of the following: next generation sequencing / systems biology / stem cell based approaches. At KCL, they will direct a 4* research programme in genetic/genomic research and contribute to undergraduate and graduate teaching.

Applicants must hold a PhD degree. Starting salary will be dependent on level of appointment. Available positions will be permanent posts.

For an informal discussion please contact Professor Gillian Bates on 020-7188-3722 or gillian.bates@kcl.ac.uk

For an information pack, please e-mail
Ms Penny Mander (gen-recruit@genetics.kcl.ac.uk)
quoting reference GB/MMG/faculty2012.

Please send a CV and summary of research interests to:
Ms Penelope Mander, Divisional Office Manager,
Department of Medical & Molecular Genetics, 8th Floor,
Tower Wing, Guy's Hospital, London SE1 9RT.
Email: gen-recruit@genetics.kcl.ac.uk
Please quote reference number.

U230020R

University of Essex

Professor in Social and Behavioural Genomics

(Ref. PR792) We are seeking to appoint a Professor to establish a new research group in Social and Behavioural Genomics. In addition to providing strategic and innovative leadership in research and teaching within the School of Biological Sciences, the successful candidate will be expected to take a leading role in establishing this new cross faculty area of research. We are looking to appoint an outstanding candidate with an international reputation and an excellent publication record.

Salary: commensurate with experience and achievements
Closing date: 25 November 2012

Apply online. If you have a disability and would like information in a different format telephone (01206) 873521/874588.

POSITIVE ABOUT DISABLED PEOPLE

http://jobs.essex.ac.uk

U230306R

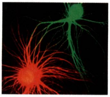
Bioinformatics/Biology Positions in Comparative & Marine Genomics

UF UNIVERSITY of FLORIDA

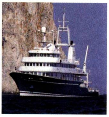
to Study Biodiversity, Origins of Complex Brains, Memory & Intelligence

Advanced Research Positions are available immediately for computer scientists, molecular and computational biologists to study Marine Biodiversity as well as Origin and Evolution of Animals, Complex Brains, Learning and Memory mechanisms. We will sequence and annotate novel genomes from a diversity of tropical and subtropical marine organisms with different levels of neural complexity, as well as identify novel biologically active compounds for the needs of biomedicine. Previous experience in genome assemblies and analysis of next generation sequencing (NGS) data such as RNA-seq, ChiP-seq, or system biology approaches to reconstruct gene regulatory circuits is preferable.

These are ideal positions for motivated, self-driven, and adventurous PhD-level scientists who would like to be integrated into interdisciplinary programs, and interact with both marine biologists and experimental neuroscientists from a location that provides excellent access to Florida's oceans. The job includes unprecedented opportunities for World Ocean traveling including active participation in scientific expeditions to study diverse tropical marine habitats (as well as scuba diving for those interested in animal collection).



Preferable Skills & Qualifications: Ideal candidates should have **1)** a Ph.D. in Bioinformatics/Genomics or Computer Science or related fields. **2)** Experience in genome and/or transcriptome assembly software, and/or mathematical modeling or epigenetic analysis or phylogenomics. **3)** Strong skills in at least one programming language (Python, JAVA or C/C++, MATLAB, R, PERL), and/or large-scale data analysis.



Contact: Prof. Leonid L. Moroz (moroz@whitney.ufl.edu); Brain & Genetic Institutes, Dept of Neuroscience and Whitney Lab for Marine Biosciences, Univ. of Florida. Please provide your CV, names of three references and evidence of your experience in the field of bioinformatics and/or computational biology and/or computer sciences. A proficiency in database development and maintenance, together with an ability to write custom scripts and/or algorithm development for informatics/visualization is a plus. Particularly helpful would be administration experience in a Linux environment and/or experience in developing multi-threaded programs or programming for a cluster environment as well as knowledge of HTMT, CSS, and JavaScript. Additional information can be found at http://www.mbi.ufl.edu/ and http://www.whitney.ufl.edu/index.php/research-programs/leonid-l-moroz.

NW230119R

Tenure Track Faculty Positions
The Department of Genetics
Washington University School of Medicine, St. Louis, MO

As part of a major expansion, the Department of Genetics is now recruiting investigators at all academic levels. Individuals with expertise in high-throughput functional genomics, including single-cell analysis; computational and systems biology; development of cellular and animal models of human disease; stem cell biology; epigenomics and analysis of gene-environment interactions; or novel strategies to evaluate the role of genetic variants in disease etiology, are encouraged to apply.

These positions offer superb scientific resources in a collegial department (http://www.genetics.wustl.edu) with traditional strengths in computational biology, genome science, and classical genetics, as well as new activities focused on neurological disorders, neurodevelopment, cellular bioenergetics, and Genomic Medicine. The department houses centers that provide access to genomic analysis for research and clinical samples as well as induced pluripotent stem cell and cellular genome engineering technologies. Department faculty members have extensive interactions with clinical and preclinical departments at the University as well as with the Institute for Clinical and Translational Science, Hope Center for Neurological Disorders, Siteman Cancer Center, and Genomic and Pathology Services (GPS@WUSTL). In keeping with the interdisciplinary nature of modern Genetics, there is the potential for joint appointments with other departments including, Medicine, Pathology & Immunology, Neurology, OB/GYN, Psychiatry, Chemistry and Computer Science & Engineering.

Applicants should submit a cover letter, curriculum vitae, a one page summary of research accomplishments and a one page description of future research plans. These documents should be submitted in electronic form in one PDF file and sent to: Jeffrey Milbrandt, Head of Department of Genetics at Genesearch2013@genetics.wustl.edu. Applicants should also arrange for three letters of reference to be sent to Genesearch2013@genetics.wustl.edu. The application deadline is December 3, 2012.

Completed applications will be reviewed on a rolling basis and interviews of qualified candidates will commence immediately and continue until the positions are filled.

Washington University is an equal opportunity affirmative action employer.

Washington University in St. Louis

NW230217R

KingsPath
Clinical Diagnostic Pathology Services
King's College Hospital NHS
NHS Foundation Trust

GSTS
Pathology

Director of Genetics Laboratories

Salary Range: £83,296 to £103,695 inclusive pa
(International Equivalent to NHS Band 9)
Ref: 196-LIS0790

This is an exciting opportunity for a Consultant Clinical Scientist in Molecular Genetics, qualified to FRCPath level (or equivalent) to lead the genetics laboratories at Guy's Hospital into the new era of diagnostic genetics and genomics. The post-holder will direct the CPA (UK) accredited Biochemical Genetics, Cytogenetics, and Molecular Genetics laboratories within the Regional Genetics Centre, and additional specialised laboratories offering genetic tests, including Pre-implantation Genetic Diagnosis and Molecular Oncology. Particular emphasis will be placed on integrating work practices and staff across these laboratories and transforming our technology to embed next generation sequencing across pathology services.

The post-holder will be employed by Guy's and St Thomas' NHS Foundation Trust and seconded to Kings Path/GSTS Pathology. The laboratories form part of the Regional Genetics Centre of the King's Health Partners Academic Health Sciences Centre (a collaboration between King's College London and Guy's and St Thomas', King's College Hospital and South London and Maudsley NHS Foundation Trusts) and have critical research and development links with the Biomedical Research Centre. You will have a vision for collaborating with colleagues from all of these organisations, commissioners and with service users, to ensure the laboratories provide an innovative, high quality and efficient clinical diagnostic service delivering maximum patient benefit. **Translational research, using next generation sequencing and bioinformatics resources, will form a main focus of the role.**

The successful candidate will have extensive experience of running genetics laboratories, including financial responsibility, and will be used to leading teams and managing change. An Honorary Chair position within Kings College London will be available to appropriately qualified candidates able to demonstrate their expert knowledge and experience through published research in peer reviewed journals.

For an informal discussion please contact Mr Adrian O'Keeffe, General Manager (adrian.okeeffe@gsts.com), Dr Shehla Mohammed, Clinical Lead for Genetics (shehla.mohammed@gstt.nhs.uk) or Professor Sean Whittaker, Clinical Director GRIDA/Joint CAG lead (sean.whittaker@kcl.ac.uk).

Interview will be mid/late December 2012.

To apply for this job please visit the Guy's & St Thomas Foundation Trust jobs page using the following link: http://jobs.gstt.nhs.uk/job/v266627

All our vacancies can be viewed on our website.

Please note we are unable to accept any applications which are not accompanied by a fully completed application form.

Our excellent benefits include final salary pension scheme. Please visit our website to find out more. Equality of opportunity is our Policy.

Guy's and St Thomas' NHS
NHS Foundation Trust

Stonewall DIVERSITY CHAMPION
INVESTORS IN PEOPLE
POSITIVE ABOUT DISABLED PEOPLE

U230308R

Wellcome Trust Genome Campus

Advanced Courses and *Scientific* Conferences

Scientific Conferences

Biomarkers for Brain Disorders
3–5 February 2013

Genomic Disorders: From 60 Years of DNA to Human Genomes in the Clinic
10–13 April 2013

Cell Senescence in Cancer and Ageing NEW
20–23 July 2013

The Leena Peltonen School of Human Genomics
18–22 August 2013

The Genomics of Common Diseases
7–10 September 2013 (Oxford, UK)

Cancer Pharmacogenomics and Targeted Therapies NEW
15–18 September 2013

Mitochondrial Disease: Translating Biology into New Treatments NEW
2–4 October 2013

Infectious Disease Genomics and Global Health
16–18 October 2013

Epigenomics of Common Diseases
7–10 November 2013

Functional Genomics and Systems Biology
21–23 November 2013

Advanced Courses and Workshops

Computational Molecular Evolution
29 April–10 May 2013

Molecular Basis of Bacterial Infection: Contemporary Research Approaches
12–18 May 2013

Functional Genomics and Systems Biology
19–28 June 2013

In silico Systems Biology
25–29 June 2013

Genetic Analysis of Multifactorial Diseases
24–30 July 2013

Exome Sequencing
8–16 September 2013

Genomic Epidemiology in Africa
15–20 September 2013 (Entebbe, Uganda)

Design and Analysis of Genetic-based Association Studies
23–27 September 2013

Genome-wide Approaches with Fission Yeast
7–14 December 2013

www.wellcome.ac.uk/hinxton

wellcome trust

U225058E

naturejobs

European Head Office, London naturejobs@nature.com
The Macmillan Building, 4 Crinan Street, London NI 9XW, UK
Tel: +44 (0) 20 7843 4961 Fax: +44 (0) 20 7843 4996

Global Head:
Nils Moeller,+44 (0) 20 7843 4953 n.moeller@nature.com
Southwest UK/RoW:
Emma Hare, +44 (0) 20 7843 4944 e.hare@nature.com
Northeast UK/Ireland:
Matthew Ward, +44 (0) 20 7014 4059 m.ward@nature.com
France/Switzerland/Belgium:
Julie Aramburu, +44 (0) 20 7843 4994 j.aramburu@nature.com
Scandinavia/Spain/Portugal/Italy:
Megan Ball, +44 (0) 20 7843 4973 m.ball@nature.com
North Germany/The Netherlands/Eastern Europe:
Kristin Doelling, +44 (0) 20 7843 4970 k.doelling@nature.com
South Germany/Austria:
Maria Kubalova, +44 (0) 20 7014 4084 m.kubalova@nature.com
Natureevents:
Matthew Clare, +44 (0) 20 7014 4015 m.clare@nature.com
Business Development Manager, Middle East:
Jon Giuliani, +44 (0) 20 7843 4922 j.giuliani@nature.com

USA Head Office, New York naturejobs@us.nature.com
75 Varick Street, 9th Floor, New York, NY 10013-1917
Tel: +1 800 989 7718 Fax: +1 800 989 7103
US Sales Manager:
Ken Finnegan, +1 212 726 9248 k.finnegan@us.nature.com

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751, Fax: +81 3 3267 8746
Asia-Pacific Sales & Business Development Manager:
Yuki Fujiwara, +81-3-3267-8765 y.fujiwara@natureasia.com
Business Development Manager, Greater China/Singapore:
Gloria To, +852 2811 7191 g.to@natureasia.com

India
3a, 4th Floor, DLF Corporate Park, Gurgaon, 122 002, India
Advertisement Sales:
Vikas Chawla, +91 1242881057 v.chawla@nature.com

Faculty positions in Bionanoscience

- Synthetic Biology
- Super-resolution Microscopy
- Cell Biology

Level Full, Associate or Assistant Professor
Term of contract Tenured or Tenure-track

The expanding Department of Bionanoscience, part of the Kavli Institute of Nanoscience Delft (www.kavli.nl), is dedicated to world-leading research at the interface between nanophysics and cell biology. This international department, established in 2010, started with three scientists with an excellent track record in single-molecule biophysics and has expanded within the last two years to 11 faculty members with a significant widening of scope towards biology. The department has excellent general facilities, including a central microscopy facility, extensive general biology facilities and a large cleanroom facility, Kavli Nanolab Delft.

We invite applications for faculty positions, from junior to senior level, in the entire range of research topics from nanotechnology to cell biology. We are particularly eager to expand in the following research areas: **synthetic biology, super-resolution microscopy (nanoscopy), and cell biology.** For the overall balance in the department, we especially welcome applications from **senior scientists and female scientists**.

We look for group leaders with an outstanding scientific track record in bionanosciences, people with a pioneering mentality who appreciate scientific and engineering challenges and who through their research, management skills, and teaching qualities can inspire co-workers and students alike. Successful candidates will be expected to establish a strong, externally funded research program and be committed to excellence in teaching at the undergraduate and graduate levels.

Please apply before December 1st, 2012. For more information about the department: **www.bn.tudelft.nl**.

For more information, please visit **www.jobsindelft.com**

TU Delft Delft University of Technology

NW230171R

Challenge the future

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Faculty Position in Materials Science and Engineering

at the Ecole polytechnique fédérale de Lausanne (EPFL)

The School of Engineering of EPFL invites applications from candidates for a **tenure-track assistant professor** within its Institute of Materials. We seek exceptional individuals who will develop and drive a research program at the forefront of the discipline, who have a strong dedication to teaching at the undergraduate and graduate levels, and who will be proactive members of a vibrant Materials community.

Top-level applications covering all areas of materials science and engineering are invited with the following topics being of particular interest:

(i) electrochemistry as applied to the processing, microstructural tailoring or performance of materials,
(ii) the processing, microstructural tailoring, characterization or performance of materials having surface functionality, and
(iii) modeling and simulation as relevant to the processing, microstructural development and performance of materials.

Start-up resources and state-of-the-art research infrastructure will be available. Salaries and benefits are internationally competitive.

The EPFL, located in Lausanne, Switzerland, is a dynamically growing and well-funded institution fostering excellence and diversity. It has a highly international campus at an exceptionally attractive location boasting first-class infrastructure. As a technical university covering essentially the entire palette of engineering and science, EPFL offers a fertile environment for research cooperation between different disciplines. The EPFL environment is multi-lingual and multi-cultural, with English often serving as a common interface.

Applications should include a cover letter with a statement of motivation, curriculum vitae, list of publications and patents, concise statement of research and teaching interests, and the names and addresses of at least five referees. Applications must be uploaded in PDF format to the recruitment web site: **imx-search13.epfl.ch**

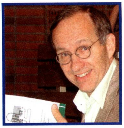
Formal evaluation of candidates will begin on **14 December 2012.**

Enquiries may be addressed to:
Prof. Harm-Anton Klok
Search Chairman
e-mail: **imx-search@epfl.ch**

For additional information on EPFL, please consult the web sites: **www.epfl.ch, sti.epfl.ch** and **imx.epfl.ch.**

EPFL is committed to increasing the diversity of its faculty, and strongly encourages women to apply.

W229323R



:insideview

profile feature

Thue W. Schwarz, *Professor of Molecular Pharmacology, Department of Neuroscience and Pharmacology, and Research Director at the Novo Nordisk Foundation Center for Basic Metabolic Research*

At a time when Danish design, food culture and drama have never been more popular, it seems all eyes are on Denmark. But does that go for science too?

Thue W. Schwartz is a professor at the Faculty of Health and Medical Sciences at the University of Copenhagen, with a background spanning Pharma, Biotech and venture capital. He tells Nature how and why Denmark's increasingly international research environment is acting as a magnet for both young and established scientific talent.

Q: What is attracting life science researchers to the University of Copenhagen?

For scientists, this is a better place than ever to live and work, the collaborative academic research environment is really super and we have a strong tradition of working closely with both local and international industry. In addition to that, we now enjoy a funding stability that not many places have.

A very important recent change was that the Novo Nordisk Foundation began funding a cluster of four big centers one by one, in addition to their traditional model of individual grants. There's also a positive interchange between the privately funded research centers and the government funded centers in fields like glycobiology, stem cells and bio-nano science. The university itself actively encourages cross-disciplinary work and is serious about things like core facilities and working infrastructure. Last but not least, Copenhagen offers an outstanding quality of life.

Q: What has the Novo Nordisk Foundation's increased support for the university achieved?

It's extraordinarily important. The first center, The Center for Protein Research, was funded in 2007. Our Center for Basic Metabolic Research followed in 2010 on a ten-year grant of 885 million DKK (approx $150 million). Since then, two more centers have been established in Stem Cell Research and Biosustainability. Although the Novo Nordisk name is in the title, the funding is unconditional; there's only one requirement and that's that we do great basic science. Having long-term funding means we can be really ambitious in terms of research plans and the recruitment of talented group leaders. They, in turn, have a very attractive environment for recruiting postdocs and PhD students internationally.

Q: How have these large centers shaped your institution?

The creation of the centers has been a hugely significant step. From the point of view of the university and the scientists, there's so much to be gained by the researchers also participating in university life and teaching students. We can create research-based courses, which are a great way of "recruiting from the bottom" and identifying the best potential PhD students. The centers have also helped us to build state-of-the-art core facilities like the Rodent Metabolic Phenotyping Center.

Alongside the special relationship with the Novo Nordisk Foundation, other private foundations have come forward with funding, including the Lundbeck Foundation and notably the Nordea Foundation's support of the Center for Healthy Ageing. Other support has come from government such as the interdisciplinary UNIK grant for Food, Fitness and Pharma even reaching social sciences and humanities.

Q: Has the creation of the centers made the university more outward looking?

Definitely—the stability of the funding means we can have a long-term strategic view. In my own Center for Basic Metabolic Research for example, we have dedicated part of the budget to creating special international research alliances. This enables the individual research director to initiate and fund joint projects with highly committed collaborators at key institutions around the world. The fact that money can be accessed rapidly and at the discretion of the research directors is unique. It has been very successful and we've already created partnerships with, for example, Rockefeller University in New York and the University of Texas Southwestern Medical Center in Dallas.

Q: What about international recruitment, what attracts people to the university?

We've been enjoying a successful period with international recruitment and I'm confident the best is yet to come. The international research alliances make a very positive contribution in flying the flag for the University of Copenhagen. They involve both young and senior scientists who are quick to grasp what a unique environment this is for a researcher.

People are also attracted to good infrastructure. The Panum Institute, a very large building that dates back to the 1970s, has seen extensive refurbishments to all lab spaces thanks to a funding partnership between the government and the university. Working here also brings some pleasant perks, like the new exercise center which has a spectacular view of Copenhagen and across to Sweden.

> "For scientists, this is a better place than ever to live and work, the collaborative academic research environment is really super and we have a strong tradition of working closely with both local and international industry"

Q: What's next?

The next important construction project is the Mærsk Building, which is where many of the internationally recruited scientists will end up. The government, university and the A. P. Møller Foundation got together to make it possible and we're immensely grateful for the generous donation of 655 million DKK (approx. $115 million) from the A. P. Møller Foundation. Speaking for myself and colleagues at the Center for Basic Metabolic Research, we're really looking forward to moving there when it's completed in 2015. Faculty and staff have had lots of input into the building's interior design and it will feature spaces designed to facilitate interaction between research groups. There'll be joint coffee areas, study spaces, even table football between open plan floors. Who knows what great scientific developments could begin with two people from opposite ends of the earth meeting at a water cooler in Copenhagen!

ADVERTISER RETAINS SOLE RESPONSIBILITY FOR CONTENT

FACULTY OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

Take your research to the highest level

The University of Copenhagen has been a front-runner in Health and Medical Sciences since it's foundation in 1479. Drawing its strength from those 533 years of history, that have seen groundbreaking work in diverse areas from Vitamin K to immunology, the faculty finds itself once again at the dawning of a new era.

It's the University's fastest growing faculty, attracting a growing population of international researchers who are actively encouraged to collaborate across disciplines and boundaries. The faculty features several centres such as the NNF Centre for Protein Research, Centre for Healthy Ageing and Danstem, the Danish Stem Cell Centre, that were conceived specifically to marry basic and applied research.

A new fifteen-storey building called the Mærsk Building is currently being developed in the heart of wonderful Copenhagen, adjacent to the existing Panum Institute. The Mærsk Building will bring together medical research, teaching and innovation on one site. Researchers of the highest quality are moving to Denmark and finding a uniquely stimulating research environment in a country that values basic science as well as innovation.

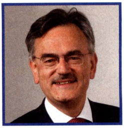
At the University of Copenhagen's Faculty of Health and Medical Sciences you don't have to choose between blue sky thinking and having your feet on the ground.

Where will your lab be?

For further information contact: Dean Ulla Wewer | dekan@sund.ku.dk | +45 3532 7483
Find out more at: http://healthsciences.ku.dk/nature

Image Credit: CF Moller, Architects

W229634R



:insideview

profile feature

Technische Universität München

Wolfgang Herrmann, *President, Technische Universität München*

"The less there is to justify a traditional custom, the harder it is to get rid of it," wrote Mark Twain. Wolfgang A. Herrmann, president of TUM (Technische Universität München) seeks to disprove Twain's notion. Herrmann is creating the first-ever U.S.-style tenure system at a German university – effectively scrapping Germany's 100+ year-old habilitation tradition. Habilitation requires would-be professors to complete the equivalent of a second PhD dissertation – and also find a faculty sponsor. Once aboard, that professor had little chance to advance within that university, and the department had no clear way to evaluate that professor. TUM is upending this tradition by creating 100 new positions by 2020 – all without the habilitation requirement, but with performance reviews after six years. Depending on the review's outcome, those professors will be given permanent status and a promotion or will leave TUM for a different job; those who achieve tenure will have an opportunity to advance further, through another evaluation process, to full professor. Herrmann tells Nature how his changes will help recruit, develop and retain top international scientific talent.

Q: In your booklet introducing TUM's tenure track system, you talk about a "paradigm shift." Why the need for such a dramatic change?

The German recruitment system has traditionally been very rigid. If you enter into a certain professorship at a certain level, you have to stay at exactly that level unless you move to another university. We think that this is not a competitive system inside the university and especially not in an international setting. The German system cannot recruit the top international talent if that talent cannot move up within the same university. There is no evaluation process in the current German system.

Q: What's the importance of evaluation and advancement?

We need to get a better record of people and how they perform. And professors deserve feedback. The best people welcome input on their job performance. Tenure-track scientists will be motivated because if they perform well, they can move up. Why would you move from a postdoc at Stanford to TUM if you have no future beyond the offered position? Young professors need an opportunity to move up. Also, why should we not promote outstanding people within our own university? In the existing system, as professors excel, their only chances of promotion are at other institutions. We want to keep the talent we develop. On the other hand, we will not tenure someone who has not qualified for promotion at the same time. The motto is "up or out."

Q: How will the tenure track system transform TUM?

We are going to create 100 new faculty positions between now and 2020 – boosting the overall number of professorships by more than 20 percent. In addition, we will transform 30 percent of the existing professorships into tenure track positions as they become vacant. We expect our reforms will make our faculty younger, more international, dynamic and competitive. We expect to fill 40 percent of our new faculty positions with people from outside Germany. We are confident of such interest, because our advancement system provides opportunities that no other German university can provide.

Q: Why are you emphasizing youth?

Right now, the average age of first professorship in Germany is around 40, with a disproportionate amount of faculty significantly older, especially when they join us from industry. We need a better mix of young, highly active people along with more experienced people to create a more dynamic intellectual environment. Along with the "young brain," we want people with international experience – someone who has had at least a year abroad.

Q: What are you doing to increase the number of female faculty?

When I started as president in 1997, we had 3 female professors out of 420. Now we are at 14 percent. For a technical university this is a lot, but it is still not enough. Our target is 25 percent by 2025. We have introduced a number of family-friendly and gender-compatible conditions, like flexible options for parental leave. We don't like quotas. But we also don't like passing over talent. Right now, far above 50 percent of Germany's medical students are female, but there are very few female medical professors and hardly any women holding key chair positions or heading departments. So, clearly female talent is out there. We just haven't done a great job of recognizing, recruiting and advancing it. Our reforms are designed to change that.

Q: How do you think other German universities will respond to your reforms?

For a couple of years there has been a strong feeling in Germany that the traditional recruitment system has to change. But no university was prepared to propose a true end-to-end tenure track system. I think we will get a lot of interest in replicating our design. But we have a big head start. We started thinking about these changes in 2006, when we competed and won in the German "Excellence Initiative." We proposed implementing these strategies in 2012, when we applied for – and won – another round of the Excellence Initiative. That "excellence" label is more important than the €2.7 billion the program apportions among the winning universities over a five-year period. It is an endorsement of our strategy. As such, we expect other universities to take notice of it.

Q: How did your own career inform the changes you are making at TUM?

I was exposed to the idea of being internationally competitive early on, since I studied under Nobel laureate E. O. Fischer in Munich. Then as a postdoc at Penn State University, I got exposed to the U.S. tenure track system. That further opened my eyes to the idea of competition. Then, during my only sabbatical, I spent time at Hoechst Corporate Research. That let me observe the career paths of industrial chemists across the world and made me wonder why academic chemists couldn't advance inside their own institution as well. Finally, I have had a number of outstanding international postdocs – none of whom stayed in Germany, because their advancement opportunities were so limited.

Q: How do you feel about changing traditions?

We have to use our traditions where they give us specific strengths. But we have to look outside for new ideas as well. The U.S.-style tenure track system is something many other countries have adopted. It is time Germany tries it as well.

ADVERTISER RETAINS SOLE RESPONSIBILITY FOR CONTENT

Opportunities for Talents

TUM

Technische Universität München

TUM is the first university in Germany to reinforce its recruitment policy by a comprehensive tenure track system. Based on best international standards and transparent performance criteria, TUM FACULTY TENURE TRACK offers quality-oriented academic career options for high-potential young scientists, from the appointment as Assistant Professor through a permanent position as Associate Professor and on to Full Professor.

The **TUM Department of Informatics** invites applications for a

Tenure Track Assistant Professorship
COMPUTATIONAL BIOLOGY

We are looking for an excellent junior scientist with a high potential for developing an internationally recognized research agenda in the field of Computational Biology. The successful candidate is expected to conduct research and teaching in Computational Biology/Bioinformatics. Applicants will have a substantial track record of comprehensive knowledge and experience in this field and will have developed methods in collaboration with experimental biologists. Important research contribution in computational biology are expected, for example for methods that support high-throughput experiments (NextGen Sequencing, Proteomics, Genomics, personalized health, and studies about variations in human and related species), methods based on machine learning, data mining, algorithmic bioinformatics, algorithmic optimization, and the optimization of complex systems in computational biology.

The scholar will contribute to teaching on the Bachelor and Master level of the Bioinformatics study section realized by BioInformatics Munich, a network bringing together the expertise in bioinformatics at several research institutions in Munich (TUM, LMU, GSF, MPI). Additional contributions to the study sections of the Department of Informatics are also expected. The active desire for interdisciplinary work within the different departments and centers of the TUM and its clinics are an essential criterion for the selection.

Applications should be submitted to: **Dean of Department of Informatics**, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany or as PDF-document to: bewerbung.dekanat@in.tum.de

The **TUM Department of Chemistry** invites applications for a

Tenure Track Assistant Professorship
COMPUTATIONAL BIOCATALYSIS

The professorship is associated with the newly established TUM Catalysis Research Center, and complements its existing competences. We are looking for an excellent junior scientist with a high potential for developing an internationally recognized research agenda in the field of Computational Biocatalysis. The successful candidate is expected to conduct research focused on the modeling of the dynamics and (catalytic) functionality of biological (macro-)molecules. The quantitative description of corresponding systems and concomitant scale-bridging processes dictates the use and development of highly efficient numerical approaches and the combination of different methodological approaches. Candidates are expected to have demonstrated excellence in research in one or several of these areas.

Applications should be submitted to: **Dean of Department of Chemistry**, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany or as PDF-document to: dekanat@ch.tum.de

The **TUM Center of Life and Food Sciences Weihenstephan** invites applications for a

Tenure Track Assistant Professorship
ECOPHYSIOLOGY OF HORTICULTURAL PLANTS

We are looking for an excellent junior scientist with a high potential for developing an internationally recognized research agenda in the field of ecophysiology of horticultural plants in protected production systems. The successful candidate is expected to conduct research and teaching in artificial light application, genotype-environment interaction, sensor-technology, resource management, or precision farming. We expect scientific objectives relevant for horticultural plant production. Desired research activities would be at the interface of plants and technics with a possible link to faculties of physics or engineering at TUM. The successful candidate will join interdisciplinary research groups of the department as well as the Hans Eisenmann-Zentrum for Agricultural Science.

Applications should be submitted to: **Dean of TUM Center of Life and Food Sciences Weihenstephan**, Technische Universität München, Alte Akademie 8, 85354 Freising-Weihenstephan, Germany or as PDF-document to: dekanat@wzw.tum.de

The **TUM Department of Informatics** invites applications for a

Tenure Track Assistant Professorship
GAMES ENGINEERING

We are looking for an excellent junior scientist with a high potential for developing an internationally recognized research agenda in the field of Games Engineering. The successful candidate is expected to conduct research and teaching in Games Engineering. The candidate should have comprehensive skills – including a university degree – in computer science. We are expecting a research record in at least one of the following fields: Realtime Rendering, Computer Animation, Visual Simulation. We are seeking for applicants who have gained experience in industrial game development, with a focus on the technical realization of immersive 3D computer games. We expect the candidate to actively take part in the Bachelor and upcoming Master program Informatics: Games Engineering, and to have an ability to pursue scientific research at international level in one of the addressed fields.

Applications should be submitted to: **Dean of Department of Informatics**, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany or as PDF-document to: bewerbung.dekanat@in.tum.de

The initial appointment will be for 6 years with an initial pay-scale grade W2. After positive evaluation in the final year, the candidate is tenured on an Associate Professor level. Tenure evaluation may be initiated after a minimum of three years only in exceptional cases, justified by extraordinarily outstanding achievements or in regard of particularly strategic significance in shaping the university's profile.

Eligible candidates have established a strong track record in the postdoctoral phase and demonstrate pedagogical and personal aptitude as well as substantial international experience. Family leave will be taken into consideration.

Supported by competitive start-up resources, candidates are expected to develop an independent and vigorous research program. Furthermore, candidates should be committed to excellence in undergraduate/graduate teaching and in supervising PhD students.

Teaching assignments include courses in the subject area and the basic courses offered by the department as well as courses for other academic TUM departments. Prerequisites for these positions are a university degree, a doctoral degree, teaching skills at university level, and additional academic achievements (according to Art. 7 and Art. 10 BayHSchPG). The ability to teach in English is a prerequisite for TUM Professors.

As part of the Excellence Initiative of the German federal and state governments, TUM has been pursuing the strategic goal of substantially increasing the diversity of its faculty. As an equal opportunity and affirmative action employer, TUM explicitly encourages nominations of and applications from women as well as from all others who would bring additional diversity dimensions to the university's research and teaching strategies. Preference will be given to disabled candidates with essentially the same qualifications. The TUM Munich Dual Career Office provides support for dual career couples and families.

Applications accompanied by supporting documentation (CV, certificates, credentials, list of publications including 3 selected reprints, list of courses taught, presentation of research and teaching strategies, third-party funding, recommendation letters from 3 leading international academics, names and addresses of at least 3 potential peer-reviewers) should be submitted by **November 15th**, 2012 to the address indicated above. All positions are to be filled as soon as possible.

W229950RL

naturejobs **PRESENTS A**

SPOTLIGHT ON THE NETHERLANDS

***NATURE* ISSUE DATE:**
15 NOVEMBER 2012

DEADLINE FOR ADVERTISERS:
9 NOVEMBER 2012

Scientific research is alive and well in The Netherlands, with total annual funding exceeding €8 billion at the end of the last decade.* Position your organization as an ideal destination for scientific excellence in the region by advertising within the *Naturejobs* Spotlight on The Netherlands. The spotlight is a prime opportunity to advertise vacancies, promote events and courses, and make important announcements to *Nature* readers alongside a dedicated editorial feature.

All job advertisements will receive a complimentary 8-week online posting on naturejobs.com and event advertisements will receive a complimentary posting on natureevents.com — the events portal for *Nature*.

*http://www.euraxess.nl/research/dutch-research-funding

For more information, contact your sales representative today -

Kristin Dölling, *Naturejobs* **T:** +44 (0)20 7843 4970 | **E:** k.doelling@nature.com

Matt Clare, *Natureevents* **T:** +44 (0)20 7014 4015 | **E:** m.clare@nature.com

naturejobs.com

nature publishing group npg

Always had a global view?

MAKE GREAT THINGS HAPPEN

Opportunities for research scientists and professionals: Looking beyond borders has been your dream since childhood days. Merck offers extraordinary global career perspectives for scientists in our Research & Development hubs in Greater Boston (USA), Greater Frankfurt (Germany), Beijing (China) or Tokio (Japan).

Contribute to challenging research projects aimed at fighting disease worldwide. Join us and take part in shaping our diversified bio pharma business. Ready for an exploration journey?

come2merck.com

MERCK

W229991R

Germany – Partner for Medical Technology

Leading-edge imaging for the bigger picture

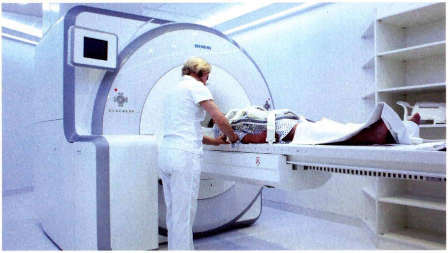
German cutting-edge research, focused support and efficient networking are helping to create the healthcare technologies of tomorrow.

Ever since Wilhelm Conrad Röntgen's discovery of X-rays in 1895, Germany has been a hotbed of innovation in medical imaging. Now the German Federal Ministry of Education and Research (BMBF) has made medical technology the focus of the most recent campaign under its "Research in Germany – Land of Ideas" initiative.

Launched at Medica 2011 in Düsseldorf, the "Germany – Partner for Medical Technology" campaign aims to raise the country's profile as a world-class location for research and development in this sector. This year the campaign will bring some of the most innovative researchers and institutions to the Radiological Society of North America (RSNA) annual meeting in Chicago, with a strong emphasis on medical imaging.

Germany is already at the centre of the global research landscape in medical technology. Scientific publications originating from Germany account for between 5 and 20 percent of the worldwide total in this area, in its various fields. Almost 9 percent of the 20 billion euros in turnover generated by the country's medical technology industry is spent on research and development. Each year around 2,300 students graduate from courses dedicated to medical technology, offered by some 20 German universities and 35 universities of applied sciences, while almost 75,000 students obtain a degree in one of the science, technology, engineering and medicine (STEM) subjects. Many of them go on to enter careers in the medical technology sector.

One of the keys to Germany's success in this sector is its network of medical technology clusters, made up of universities, research institutes, clinics, small to medium enterprises (SMEs) and large companies. Strengthening and extending these clusters is vital in order to promote and expand the country's leading role. To this end, the German Federal Government is supporting the most productive clusters in their rise to pre-eminence, with initiatives such as the leading-edge cluster competition – a flagship of Germany's High-Tech Strategy 2020.


Biograph mMR, the first clinically approved MR-PET scanner.

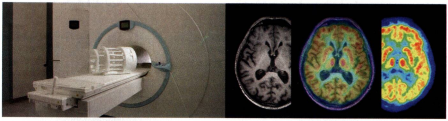
Medical imaging – a central focus

Advances in medical imaging are a central part of the progress of medical technology, delivering images with ever higher spatial and temporal resolution, and enabling ever faster and more precise diagnoses. One indicator of Germany's current strong position in this field is the fact that 50 percent of the world's magnetic resonance imaging (MRI) scanners originate here. Imaging techniques including ultrasonography, computed tomography (CT) and positron emission tomography (PET), as well as magnetic resonance tomography (MRT), continue to be developed further, and are being applied in such diverse areas as oncology, cardiovascular diseases and neurology, to name but a few.

One outstanding example is the powerful hybrid device, combining the advantages of MRI and PET, developed by scientists at the Jülich Research Centre – a member of the Helmholtz Association – in close co-operation with their colleagues at Siemens. The MR-PET scanner, applying a magnetic field of 9.4 Tesla (about 190,000 times that of the Earth), is the first instrument capable of simultaneously analysing the structure and function as well as the metabolism of the brain down to the molecular level at ultra-high field strength. In order to combine the tissue-imaging properties of MRI with PET's ability to elucidate the brain's physiological and biochemical activities, the Jülich and Siemens researchers had to overcome significant challenges, among them the incompatibility of conventional, photomultiplier-based PET detectors with magnetic fields. Instead, they employed newly developed avalanche photodiodes. Like conventional PET detectors, these measure a flow of electrons generated by photons, but in their case this flow takes place in a semiconductor layer system over distances too short to be affected by the device's powerful magnetic fields. Eventually, these technical developments enabled Siemens to build the Biograph™ mMR, the first clinically approved MR-PET scanner for simultaneous MR and PET acquisition.

While MRI measurements rely on the magnetic properties of hydrogen nuclei in tissues, magnetic particle imaging (MPI) exploits the distribution of a magnetic tracer substance in the organism. Usually, this tracer substance consists of superparamagnetic iron oxide nanoparticles. A crucial improvement of this technology – a single-sided detector – is being developed by scientists at the Institute of Medical Engineering at Lübeck University, in co-operation with their partners at Phillips Research. Whereas a conventional MPI scanner's field of view lies within a symmetric coil configuration, thereby limiting the size of specimens that can be investigated, this limit would be eliminated by using a single-sided scanner geometry. So far, resolutions of down to one millimetre in one dimension have already been achieved.

A completely new range of optical imaging methods is opened up by the multispectral opto-acoustic tomography (MSOT) technique invented by leading researchers at Helmholtz Center Munich and developed and marketed by the start-up company iThera Medical. While fluorescence imaging methods are well established in the analysis of cells in culture or of isolated tissues, their use in whole organisms had been subject to severe limitations. MSOT overcomes these limitations by using the photoacoustic effect. Tissues are irradiated with laser pulses in the near infrared part of the spectrum. The absorbed photons cause ultrasound shockwaves that can be detected and used for a reconstruction of the object being investigated. Currently this method is proving very useful for preclinical studies in animals. For example, the continuous monitoring of the development of a tumour in a mouse can help both to reduce the number of animals being used in research, and to improve the reproducibility of the results that are obtained. The range of possible future clinical applications is quite wide, from diagnosing deep-lying tissues, to high-resolution vascular diagnosis, to endoscopic applications.



UNIVERSITY AND UNIVERSITY HOSPITAL OF ERLANGEN, ERLANGEN, GERMANY

Biograph mMR 3.0 Tesla whole-body integrated MR-PET system (left) and clinical examples for simultaneously acquired MR-PET hybrid images. The Biograph mMR system combines the excellent soft tissue contrast and high spatial resolution of MR imaging with the high sensitivity of PET.

Future trends

These three examples out of many already shed some light on two of the most important trends in medical imaging today: the ever-closer convergence of classical imaging techniques like CT and MRI, and the development of optical methods with modern molecular and cell biology, which together move towards the goal of molecular imaging. These new approaches will, for example, enable physicians to reach a diagnosis earlier and to launch more focused therapies, as well as allowing them to better monitor applications such as chemotherapy. They will also help researchers gain new insights into the cause as well as the progression of diseases. The molecular imaging innovation alliance, "Innovationsallianz Molekulare Bildgebung", brings together the leading German research clusters in this field, such as the cluster MOIN CC (Molecular Imaging network), with numerous partners in the scientific community and commercial companies.

A further overarching trend is the continuous improvement and development of new software systems and algorithms for data processing and image analysis, with the goal of expanding diagnostic possibilities and enhancing the safety of therapies.

Leading BMBF-funded German clusters in this area are "Quantitative Imaging in Oncology (QUINO)", "3D Imaging in Medicine", and "German Medical Imaging in Motion (GMIM)". QUINO, which is developing new procedures in medical image analysis, is a competence network consisting of the German Cancer Research Center (Deutsches Krebsforschungszentrum; DKFZ), the University Hospital of Heidelberg and the medium-sized companies CHILI Radiology GmbH and Mint Medical GmbH.

Health technology is one of the main research and teaching topics of the Friedrich-Alexander University (FAU) in Erlangen, with its numerous spin-offs forming part of the "3D Imaging in Medicine" cluster. FAU scientists have made substantial contributions to medical imaging, telemedicine, biomaterials, and therapeutic systems. Strong links to Siemens Healthcare, Fraunhofer IIS and about 50 companies in the surrounding Nuremberg Metropolitan Region contribute to the excellent FAU research environment.

The GMIM consortium reflects the medical technology strengths of southern Germany and particularly the Munich area. It focuses both on innovative topics from different fields of preclinical, research oriented imaging, such as multimodal imaging, and on innovations in clinical imaging and image processing.

Federal Ministry
of Education
and Research

Research in
Germany
Land of Ideas

www.research-in-germany.de

The Florida State University

Strategic Faculty Recruitment in Materials for Energy Production, Conversion, Storage and Utilization

The President and the Provost of the Florida State University are pleased to announce a major interdisciplinary initiative in the area of Energy and Materials with an initial focus on materials for energy production, conversion, storage and utilization. To launch this strategic effort as many as eight tenure-track/tenured faculty positions will be filled. This faculty search is open with respect to rank and academic department. Successful candidates are expected to have a synergistic impact on existing research programs in the University's departments and interdisciplinary centers as well as develop new areas. Sustained pursuit and growth of collaborative, externally-funded research programs is an explicit goal.

Strengths at the Florida State University include energy-related materials programs in Chemistry, Physics, and Scientific Computing in the College of Arts and Sciences, and in Mechanical, Industrial & Manufacturing, Electrical & Computer, and Chemical & Biomedical Engineering in the College of Engineering. Complementing these programs are interactive centers including the National High Magnetic Field Laboratory, the Applied Superconductivity Center, the High Performance Materials Institute, the Aero-Propulsion, Mechatronics and Energy Center, and the Center for Advanced Power Systems. Linking these colleges and centers is a new Ph.D. program in Materials Science and Engineering. Robust, department-based doctoral programs in materials and related areas are also present at the University.

The Florida State University is classified as a very high research activity, doctorate-granting institution with a student population approaching 41,000. Last year the University granted over 400 doctoral degrees and had more than $220M in research expenditures. In recent years, the University has made considerable investments in research infrastructure in the sciences and engineering disciplines. The University is located in Tallahassee, the Capital of Florida, where residents have access to a broad range of cultural amenities afforded by the presence of three institutions of higher learning. The region is relatively undeveloped with close proximity to the Apalachicola National Forest, an abundance of springs, lakes and rivers as well as pristine beaches and the adjacent waters of the Gulf of Mexico.

We invite nominations for and applications from researchers active in broadly-defined areas of materials for energy production, conversion, storage and utilization including theory, computation, synthesis, fundamental materials characterization, devices and testing and proof of concept and prototype. Successful candidates will be offered highly competitive salaries and start-up packages, state-of-the-art research space and access to world-class instrumentation, computing and facilities in academic and interdisciplinary units.

Applicants are asked to provide in .pdf format a letter of application, a full curriculum vitae, the names and contact information of three professional references and a two page narrative describing their research interests that should include a clear statement as to how the candidate would complement this inter-college effort at Florida State University. Full applications must be sent electronically to **materials.search@fsu.edu**. Nominations should be sent to the same address. Review of applications and nominations will begin on December 1, 2012. Additional information about the materials programs at FSU and this faculty search can be obtained at **http://www.research.fsu.edu/materials_search/**.

The Florida State University is committed to the diversity of its faculty, staff, and students, and to sustaining a work and learning environment that is inclusive. Women, minorities, and people with disabilities are encouraged to apply. FSU is an Equal Opportunity/Access/Affirmative Action Employer.

NW229671R

SEEKING Laboratory Head: Systems Biology RIKEN Center for Developmental Biology (CDB)

The CDB plans to launch an interdisciplinary, collaborative research program in the system-level understanding of developmental and cellular phenomena. We are now seeking outstanding candidates to serve as leader of this program. The successful candidate will have expertise in either or both experimental and theoretical research, particularly in network analysis of biological functions, and will be expected to develop collaborations with research teams in the CDB. The most important evaluation criteria in this recruitment are the novelty and creativity of the research plan. The size of the laboratory space and budget will be determined based on the proposed research program.

The CDB is strongly committed to hiring and supporting staff from diverse scientific and cultural backgrounds, and welcomes applications from candidates of any gender or nationality. Candidates outside Japan are strongly encouraged to apply. Laboratories are internally funded and have access to facilities for animal care, as well as support for imaging, genomics and proteomics. The initial research term is 5 years, and is eligible for renewal for up to a single additional term pending the results of performance review.

Review of applications will begin December 1, 2012 and continue on a rolling basis until the positions are filled. Applicants should send a PDF file of the following documents in English to the e-mail address below.

1) Curriculum vitae, including a brief summary of research experience 2) Proposal for a five-year research project 3) Name and contact information of three references

Search Committee, RIKEN Center for Developmental Biology
2-2-3 Minatojima Minamimachi, Chuo-ku Kobe 650-0047, Japan
E-mail: kobe-jobs@riken.jp
URL: http://www.cdb.riken.jp/

JP229926R

SEEKING Laboratory Heads: Developmental Biology and Stem Cell Biology RIKEN Center for Developmental Biology (CDB)

Several positions are open for outstanding scientists engaged in the investigation of fundamental mechanisms of development, regeneration and stem cells, and their application to regenerative medicine. Scientists at early to middle stages of their careers will be considered. The most important evaluation criteria in this recruitment are the novelty and creativity of the research plan; those seeking to undertake new research challenges are welcomed, irrespective of their scientific background or stage of career development. The size of the laboratory space and budget will be determined based on the proposed research program and the applicant's record of achievement.

The CDB is strongly committed to hiring and supporting staff from diverse scientific and cultural backgrounds, and welcomes applications from candidates of any gender or nationality. Candidates outside Japan are strongly encouraged to apply. Laboratories are internally funded and have access to facilities for animal care, as well as support for imaging, genomics and proteomics. The initial research term is 5 years, and is eligible for renewal for up to a single additional term pending the results of performance review.

Review of applications will begin December 1, 2012 and continue on a rolling basis until the positions are filled. Applicants should send a PDF file of the following documents in English to the e-mail address below.

1) Curriculum vitae, including a brief summary of research experience 2) Proposal for a five-year research project 3) Name and contact information of three references

Search Committee, RIKEN Center for Developmental Biology
2-2-3 Minatojima Minamimachi, Chuo-ku Kobe 650-0047, Japan
E-mail: kobe-jobs@riken.jp
URL: http://www.cdb.riken.jp/

JP229925R

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

Faculty Positions in Computational Bioscience Research Center (CBRC) of the Computer, Electrical and Mathematical Sciences and Engineering Division at King Abdullah University of Science and Technology (KAUST)

The Computational Bioscience Research Center (CBRC) at King Abdullah University of Science and Technology (KAUST) invites applications for the faculty positions at all levels (Full, Associate, and Assistant Professor).

CBRC, http://cbrc.kaust.edu.sa, is part of the Computer, Electrical and Mathematical Sciences and Engineering Division (CEMSE) (http://cemse.kaust.edu.sa/). CBRC develops novel enabling bioinformatics methods and applies them in life science research. The activities of the Center are both in development of new methods and algorithms for computational biology and bioinformatics, their implementation on high performance computing platforms, and, their application in systems biology, microbiology, biotechnology and related fields. The focus of the CBRC is on modeling the key processes in the living cell so as to be able to explore the way cells react to different challenges. Bioinformatics research is motivated by and complemented with CBRC's internal and collaboratively conducted experimental work that helps refine and validate in silico-derived hypotheses potentially useful for various applications.

We are interested in applicants with the background primarily in mathematics, computer science, and engineering, though excellent candidates with life science backgrounds will be considered. The research focus of the candidates should be in one or more domains spanning computational biology, bioinformatics, systems biology, genomics, data integration, mathematical modeling, data mining, and artificial intelligence. Teaching interests may fall within the broad scope of the CEMSE division and, potentially, within Biological and Environmental Sciences and Engineering Division (BESE, http://bese.kaust.edu.sa/).

To apply for this position, please send a single PDF or word file containing: a cover letter, curriculum vitae, complete publication list, research plan, teaching statement, and contact information of four references to CBRC.recruitment@kaust.edu.sa with attention to Professor Vladimir Bajic.

Applications received by January 15, 2013 will receive full consideration and positions will remain open until filled.

About KAUST

King Abdullah University of Science and Technology (KAUST) is an international, graduate research university dedicated to advancing science and technology through interdisciplinary research, education, and innovation. Located on the shores of the Red Sea in Saudi Arabia, KAUST offers superb research facilities, generous assured research funding, and internationally competitive salaries, attracting top international faculty, scientists, engineers, and students to conduct fundamental and goal- oriented research to address the world's pressing scientific and technological challenges related to the sustainability of water, food, energy, and the environment.

RW230251R

www.kaust.edu.sa

Imperial College London

Faculty of Natural Sciences

PhD Studentships in Science and Mathematics

Imperial College London embodies and delivers world class scholarship, education and research in science, engineering, medicine and business.

The Faculty of Natural Sciences is offering over 150 PhD studentships for Home, EU and Overseas students starting in academic year 2013/14, including fully funded PhD studentships.

The Faculty of Natural Science's research strategy is to harness the strength and breadth of its research to meet the changing needs of society, industry and healthcare, and to address the global challenges of climate change, energy, global health and security.

We are seeking high calibre graduates across the full breadth of science and mathematics, both in core disciplines and at interfaces between subjects, to undertake intellectually challenging projects in a number of research areas ranging from Catalysis and Systems & Synthetic Biology through to Planetary Physics, Climate Change and Mathematical Finance.

Successful candidates will form part of a cohort of PhD students working on projects in the Departments of Chemistry, Life Sciences, Mathematics and Physics, the Grantham Institute of Climate Change and the Centre for Environmental Policy. For full details of the breadth of research areas please refer to http://www3.imperial.ac.uk/naturalsciences/research

Imperial College London provides an excellent PhD training environment that supports students in the acquisition of high-level research skills and provides a wide diversity of opportunities to participate in both core-disciplinary and multidisciplinary research. The College's world-class facilities ensure that students are exposed to the latest techniques and acquire skills that will enable them to become leaders in their chosen field. Our Graduate School offers an award winning professional development programme that supports students throughout their programme of study and research and equips them for their future careers.

Over 150 studentships are available. Further information about PhD study at Imperial College London is available at http://www3.imperial.ac.uk/registry/admissions/howtoapplypg and enquiries can be made to fonspgenquiries@imperial.ac.uk

POSITIVE ABOUT DISABLED PEOPLE

Stonewall
DIVERSITY CHAMPION

Committed to equality and valuing diversity. We are also an Athena Bronze SWAN Award winner, a Stonewall Diversity Champion and a Two Ticks Employer.

U230293R

UNIVERSITY OF CAMBRIDGE

A world of opportunities
www.cam.ac.uk/jobs/

Herchel Smith Postdoctoral Research Fellowships

Academic Division
£27,578 - £35,938 pa
Limit of tenure: Funding is available for a maximum period of three years

The Managers of the Herchel Smith Postdoctoral Research Fellowships Fund invite applications within the fields of Biological Sciences, Pure Mathematics, Biochemistry and/or Organic Chemistry & Biophysics or Geophysics.

All Fellowships are to be held from 1 October 2013 or otherwise by negotiation. Fellowships are available for between two and three years and provide an opportunity for independent research, although the holders will usually work in close collaboration with an established research group.

In accordance with Dr Smith's will, candidature is limited to candidates who have obtained their PhD degree, or equivalent, within the last three years at any university but normally excluding Cambridge and Harvard.

The stipend will be on the University's Postdoctoral Research Associate scale, currently between £27,578 to £35,938 pa, with a research allowance of £13,000 in the first year and £11,000 thereafter.

Further details are available at http://www.herchelsmith.cam.ac.uk/fellowships/ or from the Secretary to the Fund Managers, Tel. +44 (0)1223 764987, e-mail hsf@admin.cam.ac.uk

Quote Reference: AK22232. Closing date: 3 December 2012.

The University values diversity and is committed to equality of opportunity.

U230305R

---

epp **engineering and public policy**
Preparing Technical Leaders to Address Policy Issues that Involve Science and Technology

Faculty Openings in EPP at CMU

The Department of Engineering and Public Policy at Carnegie Mellon seeks the following faculty candidates:

- A candidate at any level with a strong technical background and knowledge of the relevant literatures to address problems in the management of technical innovation and R&D policy from "inside the black box."
- A senior economist with a technical background to work on problems in technology and policy and occupy the Lester B. Lave chair between EPP and the Tepper School of Business.
- An engineer at any level to address innovation in energy technology.
- Other candidates with engineering or science backgrounds and demonstrated accomplishment in technology and public policy.

See www.epp.cmu.edu.
Send resume, references, and 2-3 sample publications to A. Loucks (aloucks@andrew.cmu.edu).

NW229681R

---

naturejobs

Read our Postdoc articles and news in the **News and Comment** section of **naturejobs.com**

- Helpful advice
- Real life experiences
- Follow the highs and lows of postdoc life

Save time! Save your searches on **naturejobs.com** as a Job Alert email or RSS feed

naturejobs
making science work

Johnson & Johnson
FAMILY OF COMPANIES

What type of mark will you make?

Janssen Research & Development

Janssen Research & Development, LLC, discovers and develops some of the most innovative technology platforms and biologic medicines today. We are a world leader in monoclonal antibody technology and have a strong pipeline of novel antibodies in development.

Join our Informatics Group and be part of something revolutionary.

Imagine working at the cutting edge of quantitative biomedicine to tackle the challenges of drug discovery... surrounded by reminders of the earliest days of American history. Janssen Research & Development is based in Spring House, Pennsylvania, a quiet escape from urban life that had its beginnings in the 1700s. Within our Spring House campus, you will be interacting with a multidisciplinary team of researchers focused on bringing therapies to patients. Looking for more excitement? Philadelphia, New Hope, Princeton, Lancaster, New York City and the New Jersey beaches are an easy, scenic drive away. Can you picture yourself here? Come work with us.

We are looking for experts from around the globe with a driven, passionate, and pioneering spirit to join our team. This is an outstanding opportunity to make a difference in the lives of people living with chronic diseases, and to work in a beautiful setting that is far from ordinary.

We are currently recruiting outstanding scientific leaders with experience and expertise for the following roles:

- **Head, Computational & Systems Biology**
  Requisition# **3489120829**
- **Head, High Performance Computing**
  Requisition# **2671120821**
- **Senior Scientist, Computational & Systems Biology**
  Requisition# **3468120829**
- **Senior Scientist, High Performance Computing**
  Requisition# **2568120816**

Relocation packages are available.

BE VITAL
careers.jnj.com

Help us make history.

If you need assistance locating these positions on our website, please email JRegan14@ITS.JNJ.com.

© Johnson & Johnson Services, Inc. 2012. Janssen Research & Development, LLC, is a member of the Janssen Pharmaceutical Companies of Johnson & Johnson. Johnson & Johnson companies are equal opportunity employers.

NW229896R

UNIVERSITY OF
OXFORD

www.ox.ac.uk/jobs

Medical Sciences Division

Sir William Dunn School of Pathology

SILVER
Athena SWAN
Charter for Women in Science

University Lecturership in Cellular Pathology and Proteomics

in association with an Official Fellowship at Linacre College

Grade: 36S ULNTF: £42,883 - £57,581 per annum

Applications are invited for a University Lecturer (equivalent to Assistant Professor) to lead a new group at the Dunn School of Pathology. The primary criteria for selection will be scientific excellence and clear evidence of a commitment to teaching. The Dunn School, which has outstanding facilities and resources, has a research focus on the fundamental cell and molecular biology underlying a broad range of disease mechanisms. It also takes a lead role in teaching preclinical science to medical students.

You should have completed a period of postdoctoral training, or equivalent, with an excellent track record, and show outstanding potential for independent research in a topic related to proteomics.

More details about the post and full application instructions are available at **http://www.ox.ac.uk/about_the_university/jobs/fp/** Informal enquiries can be made to the incoming head of department, Matthew Freeman at Matthew.Freeman@path.ox.ac.uk

Further particulars, including detailed selection criteria and guidance on applying for this post, are available on the above link or from gillian.helstrom@path.ox.ac.uk

Applications, including a covering statement and full CV, and naming three referees should be received no later than 12 noon on Friday 14 December 2012. If you have a query about how to apply, please contact Gillian Helstrom at gillian.helstrom@path.ox.ac.uk or telephone 01865 275503.

Applications are particularly welcome from women and black and minority ethnic candidates, who are under-represented in academic posts in Oxford.

Committed to equality and valuing diversity

U230296R

Call for Applications

Applications are now being accepted for the 2013 NASA Space Radiation Summer School, a three-week course designed to offer graduate students, postdoctoral fellows, and faculty an integrated curriculum of radiation biology, radiation chemistry, and physics culminating in hands-on accelerator-based experiments using the synchrotron facility at the NASA Space Radiation Laboratory.

Up to 15 students will be selected for the course, tentatively scheduled for June 3–21, 2013 at the Brookhaven National Laboratory (BNL) on Long Island, New York. Topics will include DNA damage and repair, genotoxicity, cell cycle checkpoints and apoptosis, mutagenesis, genomic instability, epigenetics, cell and tissue signaling, neurodegeneration, systems biology, and the relationship of these processes to carcinogenesis and late degenerative effects following exposure to space radiation, as well as the space radiation environment, physics and biochemistry of charged particle interactions, track structure, dosimetry, accelerator operations, and space radiation protection.

Course faculty consists of leading university and national laboratory biologists and physicists actively engaged in NASA space radiation research and BNL experts in heavy ion experimentation and methods.

Application instructions are available online at

www.dsls.usra.edu/spacerad/2013/

Application Deadline - 11:59 p.m. CT, Sunday, February 10, 2013

U.S. citizens and foreign nationals may apply. All selected students must satisfy BNL/DOE safety and security requirements in order to be admitted. Selected students must also demonstrate oral and written proficiency in the English language.

Travel expenses in the U.S. and room and board will be covered for selected participants. Successful applicants from outside the U.S. must provide for their travel to/from New York/Long Island.

Course sponsors are the NASA Space Radiation Program, Pacific Northwest National Lab, U.S. Dept. of Energy, Brookhaven National Lab, and Universities Space Research Association.

NW229682R

UNIVERSITÄT BASEL

Professor in Molecular Stem Cell Biology

The Biozentrum of the University of Basel (Switzerland) invites applications for a Professorship in Molecular Stem Cell Biology at the rank of Assistant (tenure-track), Associate or Full Professor, depending on the qualifications of the applicant.

We are seeking outstanding candidates in basic molecular and cellular stem cell biology. Areas of research interests include, but are not limited to, stem cell-niche interactions, programming/reprogramming, differentiation/dedifferentiation and asymmetric cell division.

The ideal candidate should establish an internationally competitive research program, will complement currently existing strengths of the Biozentrum and will contribute to the Basel Stem Cell network. The successful candidate is expected to actively participate in teaching at the undergraduate and postgraduate level.

The Biozentrum offers an outstanding scientific environment and an attractive research endowment. Basel is the home of a vibrant international life science community and provides a high standard of living and a superb cultural atmosphere.

Applications, including CV, list of publications and a research program summary, should be sent by e-mail (pdf or zip to Prof. Dr. Jörg Schibler, Dean, Faculty of Science, University of Basel, Klingelbergstrasse 50, 4056 Basel, Switzerland, to dekanat-philnat@unibas.ch. For informal inquiries please contact Prof. Dr. Erich A. Nigg (erich.nigg@unibas.ch, phone: +41-61-267 16 56).

The deadline for receipt of applications is December 31, 2012. The University of Basel is an equal opportunity employer and encourages applications from female candidates.

UNI
BASEL

W230118R

Follow *Naturejobs* on Twitter

For the latest career news: **twitter.com/naturejobs**
For all job listings: **twitter.com/naturejobs_feed**

twitter

bioss

CENTRE FOR BIOLOGICAL
SIGNALLING STUDIES

CLUSTER OF EXCELLENCE

BIOSS Centre for Biological Signalling Studies

UNI
FREIBURG

BIOSS, the Research Centre for Biological Signalling Studies at the University of Freiburg (Germany) provides excellent research opportunities to study the mechanism of biological signalling. The central research idea of BIOSS is to initiate and promote a dialectic process between scientists using analytical (dissecting) and synthetical (rebuilding) approaches in signalling research in an interdisciplinary strategy together with various Faculties of the University of Freiburg as well as research centres like the Max Planck Institute for Immunobiology and Epigenetics. BIOSS is funded by the Excellence Initiative of the German Federal and State Governments.

Research Area Ⓒ Synthesis and Engineering

The Faculty of Biology at the University of Freiburg (Germany) and BIOSS invite applications for a

Junior Professorship (W1) with tenure track option for Protein Evolution and Engineering

We are looking for candidates with outstanding international reputation and experience in the field of protein engineering and directed protein evolution. In teaching, the candidate is expected to contribute to the curricula in Genetics and Developmental Biology, Translational Biology, Immunology and Microbiology. Close interactions and cooperation with the established research center SFB 746 ('Functional specificity by coupling and modification of proteins)' would be appreciated. For further information please visit the BIOSS homepage www.bioss.uni-freiburg.de or the homepage of the faculty of biology www.bio.uni-freiburg.de.

The successful candidate will have completed his/her doctoral studies with distinction. Employment conditions are based on § 51 Landeshochschulgesetz. The appointment is limited to four years with the possibility of extension for another two years after evaluation.

Deadline for applications is 15.12.2012. Complete applications along with all pertinent documents (exclusively in electronic form; Email: dekanat@biologie.uni-freiburg.de) should be addressed to the Dean of the Faculty of Biology, Schänzlestr. 1, D-79104 Freiburg.

Research Area Ⓐ Intracellular Signalling Pathways

The University of Freiburg (Germany) is accepting applications for the faculty position of a

Full Professorship (W3) for Pharmaceutical Technology and Biopharmacy

This chair is established by the Excellence Cluster of the University Freiburg: BIOSS Centre for Biological Signalling Studies (www.bioss.uni-freiburg.de) at the Institute of Pharmaceutical Sciences. We seek a visionary scientist with strong research credentials in "Molecular dynamics and engineering of membrane systems" or similar areas. Close cooperation within the BIOSS program and projects in pharmacy, biology or medicine are expected. Experience in teaching pharmaceutical technology and in acquisition of research grants is required.

Deadline for applications is 15.12.2012. Applications including necessary supporting documents should be sent (exclusively in electronic form; Email: kontakt@bioss.uni-freiburg.de) to: Dekan der Fakultät für Chemie, Pharmazie und Geowissenschaften, Hebelstraße 27, D-79104 Freiburg.

Research Area Ⓒ Synthesis and Engineering

The Faculty of Biology at the University of Freiburg (Germany) and BIOSS invite applications for a

Full Professorship (W3) for Integrative Signalling Research

We are looking for candidates with outstanding international reputation and experience demonstrating that he/she will be able to combine approaches from biology and medicine with physics, chemistry and engineering to gain an integrated view of signal transduction systems in cells. In teaching, it is expected that the candidate will design a concept to impart knowledge in the curricula of the B.Sc., M.Sc. and teacher training program Biology. Close interactions and cooperations with the established research centres SFB 746 ('Functional specificity by cou-pling and modification of proteins') and SFB 850 (`Control of cell motility during morphogenesis, tumour invasion and metastasis´) will be appreciated.

For further information please visit the BIOSS homepage: www.bioss.uni-freiburg.de or the homepage of the faculty of biology www.bio.uni-freiburg.de. The position requires a completed university education, including a PhD and German habilitation or equivalent scientific accom-plishments, and teaching experience.

Deadline for applications is 15.12.2012. Complete applications along with all pertinent documents (exclusively in electronic form; Email: dekanat@biologie.uni-freiburg.de) should be addressed to the Dean of the Faculty of Biology, Schänzlestr. 1, D-79104 Freiburg.

The University of Freiburg is an equal opportunity employer. Applications of women are strongly encouraged. Handicapped candidates with equivalent qualifications will be given preference. Applicants should ask for an application form for professorship at the Deans office.

www.bioss.uni-freiburg.de

W230101R

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

The 2013 Louisa Gross Horwitz Prize for Biology or Biochemistry

The Louisa Gross Horwitz Prize was established under the will of the late S. Gross Horwitz through a bequest to Columbia University and is named to honor the donor's mother. Louisa Gross Horwitz was the daughter of Dr. Samuel David Gross (1805-1889), a prominent surgeon of Philadelphia and author of the outstanding *Systems of Surgery* who served as President of the American Medical Association.

Each year since its inception in 1967, the Louisa Gross Horwitz Prize has been awarded by Columbia University for outstanding basic research in the fields of biology or biochemistry. The purpose of this award is to honor a scientific investigator or group of investigators whose contributions to knowledge in either of these fields are deemed worthy of special recognition.

The Prize consists of an honorarium and a citation which are awarded at a special presentation event. Unless otherwise recommended by the Prize Committee, the Prize is awarded annually. Dr. Richard Losick, Harvard University, Cambridge, MA, Dr. Joe Lutkenhaus, University of Kansas Medical Center, Lawrence, KA and Dr. Lucy Shapiro, Stanford University School of Medicine, Stanford, CT were the 2012 awardees.

QUALIFICATIONS FOR THE AWARD

The Prize Committee recognizes no geographical limitations. The Prize may be awarded to an individual or a group. When the Prize is awarded to a group, the honorarium will be divided among the recipients, but each member will receive a citation. Preference will be given to work done in the recent past.

Nominations must be submitted electronically at: **http://www.cumc.columbia.edu/research/horwitz-prize**
All communications and materials must be written in the English language.

Re-nomination(s) are by invitation only.

Nominations should include:
1) A summary, preferably less than 500 words, of the research on which this nomination is based.
2) A summary, preferably less than 500 words, of the significance of this research in the fields of biology or biochemistry.
3) A brief biographical sketch of the nominee, including positions held and awards received by the nominee.
4) A listing of up to ten of the nominee's most significant publications relating to the research noted under item 1.
5) A copy of the nominee's curriculum vitae.

Deadline date: January 31, 2013

NW230218R

USA UNIVERSITY OF SOUTH ALABAMA
CHILDREN'S & WOMEN'S
HOSPITAL

Alabama: Neonatologist

As part of efforts to expand its scholarly and clinical activities, the Department of Pediatrics at the University of South Alabama College of Medicine is currently seeking candidates for a faculty position at the Assistant Professor level in the Division of Neonatology. The successful candidate will interact with five full-time faculty members in the Division of Neonatology. Duties include patient care activities in the neonatal unit at the USA Children's and Women's Hospital and teaching medical students and residents.

Our Neonatal ICU is a Level III 70-bed unit, soon to be renovated and expanded to a 100 bed unit that serves as the primary referral center for the region. Unique opportunities for clinical research as well as collaborative research in the USA Center for Lung Biology are available. Experience in studies of the molecular biology of the pulmonary system is highly desirable. We offer competitive compensation and benefits. This is an outstanding opportunity to join a congenial, diverse, expanding department committed to providing its faculty with the opportunity to achieve academic and professional growth while enjoying the climate, affordability, and high quality of life of the central Gulf Coast.

For information or to apply, please contact:
David Gremse, MD, Professor and Chair
USA Department of Pediatrics
1504 Springhill Avenue, Room 5309
Mobile, AL 36604
Email: dgremse@usouthal.edu Phone: 251-434-3919

USA is an Equal Opportunity/Equal Access Employer

NW230002R

ONE TEST

And a family is reunited

Angela was 21 days old when she was kidnapped in broad daylight. Her mother, Brenda, was on her way to the doctor when two women assaulted her, snatched Angela from her arms, and took off in a car in Guatemala City.

Two months passed before Angela was found. If it were not for a DNA test—a test created by Life Technologies—there would have been no proof Angela belonged to Brenda.

The test result revealed a 99.9% positive DNA match. Finally, Angela got to go home.

Shaping Discovery.
Improving Life.

Angela and Brenda Corado

Join in to stand out. Go to **lifetechnologies.com/careers**

life technologies™

Life Technologies is an Equal Employment Opportunity/Affirmative Action employer.
©2012 Life Technologies Corporation. All rights reserved. CO03104 0712

NW229793R

Department of Health and Human Services
National Institutes of Health
National Cancer Institute
Bethesda, Maryland

NATIONAL CANCER INSTITUTE

TENURE-ELIGIBLE/TENURE-TRACK PRINCIPAL INVESTIGATORS

The Laboratory of Genome Integrity (LGI), Center for Cancer Research (CCR), National Cancer Institute (NCI), is accepting applications to fill Principal Investigator positions at the Tenure-Track or Tenured levels. The laboratory focuses on understanding the pathways that maintain genomic integrity, the intersection of these pathways with normal cellular physiology and cancer, as well as the application of these insights to translational research. The LGI is led by Dr. Andre Nussenzweig and is located on the campus of the NIH in Bethesda, MD. To learn more about the laboratory, please visit the LGI s website: http://ccr.cancer.gov/staff/staff.asp?Name=nussenzweig.

We are seeking candidates whose research programs address basic biological problems and who have demonstrated excellence, originality, and productivity in research in the areas of biochemistry, and cellular and molecular biology. We are especially interested in applicants using innovative approaches to solve key questions in the areas of aging, DNA repair, chromosome and/or genome function.

Candidates must have a Ph.D., or M.D. degree or equivalent doctoral degree in a relevant field with post-doctoral experience and a proven ability to conduct innovative research. The incumbent will receive research support for developing a state-of-the-art laboratory that includes sufficient space, equipment, and supply budget in order to sustain a research program. Salary is commensurate with research experience and accomplishments.

Interested applicants should submit a cover letter, curriculum vitae including a list of publications, a brief statement of research accomplishments and future plans, and request three letters of recommendation to be sent electronically to christy.worch@nih.gov. Review of applications will begin on or about by January 15, 2013.

Applications will be accepted until the position is filled. This position is not restricted to U.S. citizens only.

DHHS and NIH are Equal Opportunity Employers

NW230222R

Georgia Tech College of Engineering

EMORY UNIVERSITY
SCHOOL OF MEDICINE

Faculty Position Openings
Atlanta, Georgia

Georgia Tech Campus

The Wallace H. Coulter Department of Biomedical Engineering, a joint department between Georgia Tech's College of Engineering and Emory University's School of Medicine, invites applications for tenure track faculty at the assistant, associate, and full professor levels. Applicant's research should be synergistic with the existing strengths in the department, which are described under Research Overviews on our website, **www.bme.gatech.edu.** We are particularly interested in applicants with expertise in cardiovascular biology/bioengineering, systems biology/synthetic biology, image processing/analysis, and pediatric bioengineering. However, outstanding applicants in other areas of research will also be considered.

Applicants must hold a doctoral degree (PhD, MD or MD/PhD) in engineering, the natural sciences, or medicine. Successful applicants will be expected to develop a funded research program and to participate in teaching and advising undergraduate and graduate students.

Applications with a statement of research and teaching interests, curriculum vitae, and names and complete contact information for three professional references should be sent to: Amy Tang, Department of Biomedical Engineering, Georgia Tech and Emory University, 313 Ferst Drive, N.W., Atlanta, GA 30332 (Phone: 404-385-7177; e-mail: amy.tang@bme.gatech.edu). Applicants may also apply on-line at: **https://www.bme.gatech.edu/jobs/**

Application deadline: January 31, 2013

Applications from women and underrepresented minorities are encouraged.

Georgia Institute of Technology and Emory University are AA/EOE employers.

NW230262R

REGENERON
science to medicine™

POST-DOCTORAL TRAINING PROGRAM OPPORTUNITY

Learn from a leader. Be part of the next life-changing innovation.

Discover a career at the forefront of new discoveries with a top ten biotechnology company focused on changing lives through groundbreaking science. Regeneron's post-doctoral training program will give you the chance to be surrounded by brilliant minds and place cutting-edge resources and technology for your research at your fingertips.

Launching in 2013, our post-doctoral program offers three years of formal post-doctoral training directed by an award-winning educator, frequent learning opportunities including journal clubs and data dumps, and the ability to pursue your passion for science in a driven environment.
Learn more at **www.regeneron.com/scienceeducation_professional**

Regeneron is an equal opportunity employer. As such, applicants are considered for employment without regard to race, color, religion, sex, national origin, age, sexual orientation, marital status, disability, genetic information, military or veterans status, or any other category protected under applicable federal, state or local law.

NW229821R

Department of Health and Human Services, National Institutes of Health, National Cancer Institute, Center for Cancer Research

Tenure-Track or Tenure-Eligible Physician-Scientist, Dermatology Branch

Job Description

The Dermatology Branch, Center for Cancer Research, National Cancer Institute is a national leader in investigative dermatology, cutaneous biology and cutaneous immunology. We are seeking an outstanding tenure-track or tenure-eligible physician-scientist to carry out original, independent, and innovative laboratory studies in a stimulating research environment.

About NCI's Center for Cancer Research

The Center for Cancer Research (CCR) is the largest intramural research component of the National Cancer Institute (NCI). The CCR provides extensive opportunities for collaboration, and allows scientists and clinicians to undertake high-risk, high-impact laboratory- and clinic-based investigations. CCR's enabling infrastructure also facilitates clinical studies at the NIH Clinical Center, the world's largest dedicated clinical research complex. Investigators are supported by a wide array of intellectual, technological and research resources, including animal facilities and dedicated, high-quality technology cores in areas such as imaging/microscopy, chemistry/purification, mass spectrometry, flow cytometry, genomics/DNA sequencing, transgenics and knockout mice, arrays/molecular profiling, and human genetics/bioinformatics. For an overview of CCR, visit http://ccr.cancer.gov/.

Location

Dermatology Branch laboratories are located on the campus of the National Institutes of Health in Bethesda, Maryland, a suburb of Washington, D.C.

Eligibility/Salary

Applicants must have an M.D. degree and be Board-certified in Dermatology. The successful candidate will have expertise and a record of achievement in an interesting area of contemporary biomedical research that is relevant to important diseases such as cancer. A clinical dimension that complements a laboratory research program is desirable, and a commitment to training and mentoring is required. The incumbent will direct an independent research program consisting of post-doctoral fellows and support personnel funded by the NCI intramural research program. Research support for developing a state-of-the-art laboratory that includes sufficient space, equipment and a supply budget that will sustain a productive research program will be provided.

Salary is competitive and commensurate with research experience and accomplishments, and a full Civil Service package of benefits (including health insurance, life insurance, retirement and a Thrift Savings Plan) is available. Candidates may be eligible for the NIH Loan Repayment Program (http://www.LRP.NIH.gov).

Application

Interested individuals should submit a letter of interest, curriculum vitae, brief statement of research interests and future plans (1-2 pages) electronically to Ms. Sallie Baird (sbaird@mail.nih.gov) or by mail to the address below. Applications will be reviewed beginning on > 60 days after the first public posting and will be accepted until the search process has been completed.

Dr. Daniel Fowler
Search Committee Chair
Dermatology Branch
Center for Cancer Research, NCI
c/o Ms. Sallie Baird
9000 Rockville Pike
Building 10, Room 12N238 MSC 1908
Bethesda, MD 20892-1908

DHHS, NIH and NCI are Equal Opportunity Employers. The NIH is dedicated to building a diverse community in its training and employment programs.

NW230050R

Université de Montréal

Université de Montréal

Faculty of Medicine – Department of Biochemistry

Research Assistant Professor position in Biochemistry

The Department of Biochemistry is at the core of a research-intensive environment in biomedical sciences at the Université de Montréal, one of the top three research universities in Canada. The Department currently has 31 Professors and 41 accredited members in the areas of Bioinformatics, Cell signaling, Clinical Biochemistry, Gene expression, Genomics, Molecular Evolution, Functional Proteomics and Structural Biology. The members of the department form a dynamic and interactive group, our graduate programs are home to more than 200 students and our faculty members regularly publish in high-impact journals and are well funded.

We will consider strong candidates in the area of **Structural Biology of Macromolecules**. We encourage candidates with an expertise in X-ray crystallography, using *in silico* and/or experimental approaches to molecular recognition and design of bioactive small molecules or of therapeutic macromolecules. Expertise with membrane proteins is a plus, and the successful candidate could benefit from participation in the membrane protein research group (GÉPROM) at the Université de Montréal.

Candidates are expected to develop an internationally competitive research program, to participate in teaching at the graduate and undergraduate levels and to contribute to departmental life and administration. Candidates will have a Ph.D. in Biochemistry or in a related discipline, experience in research at the post-doctoral level and an excellent publication record. The language of instruction at the University of Montréal is French; language training is offered to non-French speakers who will be required to be sufficiently proficient to teach in French after a period of three years.

The Université de Montréal offers competitive salary and benefits packages according to the collective agreement. The starting date is flexible and could be as early as fall 2013. Applications will be accepted until **November 30st, 2012**. Interested candidates should send a letter of interest, complete CV, research plan and the contact information for three referees, in electronic format, to:

Dr. Christian Baron, Chair, Department of Biochemistry, Université de Montréal

Email: christian.baron@umontreal.ca

The full text of the ad can be found at: http://www.biochimie.umontreal.ca/poste_stages/autres_EN.html

NW230009R

Research Scientist (Molecular Catalysis)

Division: Chemical Sciences

The Joint Center for Artificial Photosynthesis (JCAP) will focus on the development of molecularly designed, inorganic catalysts for electrocatalytic water splitting. JCAP is a Solar Fuels Innovation Hub recently funded by the Department of Energy (122 M, 5 years) with physical location at the sites of its major partners, Caltech (South) and Lawrence Berkeley National Laboratory (North). The researcher will take major responsibility for the synthetic and catalyst evaluation efforts in the Molecular Catalysis Project located at JCAP North Site.

For a Molecular Catalysis Chemist, this position offers a unique opportunity for an early career researcher to assume a key role at the inception of the JCAP research effort in development of efficient electrocatalysts for artificial photosynthesis. Anticipated activities involve novel synthetic approaches, including those based on molecular design and precursor chemistry, for targeting novel catalytic structures such as inorganic clusters, nanostructures, and stabilized, surface-bound catalytic centers. Significant effort will also involve electrochemical studies for the evaluation of new catalysts, and for characterizations of active catalytic centers and catalytic mechanisms.

For more information about this position
and to apply online, visit
http://go.lbl.gov/75229
and follow the instructions to complete
the application process.

Berkeley Lab is an affirmative action/equal opportunity employer committed to the development of a diverse workforce.

BERKELEY LAB
Lawrence Berkeley National Laboratory

NW230216F

Would you like to contribute to innovative research with the goal of improving human health? **The Center for Proteomic Chemistry (CPC) within the Novartis Institutes for BioMedical Research (NIBR)** has a number of exciting opportunities for highly motivated and talented scientists to join a global team to develop and apply structural and biophysical tools in the study of protein-ligand complexes.

Research Investigators – Structural Biology. Use your scientific expertise to advance our structure-based drug design efforts in drug discovery. Your task will be the determination of high-resolution, three-dimensional crystal structures of medically important target-ligand complexes.

We are seeking two protein crystallographers to join our structural biophysics groups in Cambridge and Basel. Experience in conducting experiments related to all aspects of crystallography is required. A track record of scientific achievements in working with integral membrane proteins is highly preferred.

For details on these lab head positions and to apply, visit **http://www.novartis.com/careers/**. For the position in Cambridge, see Job Code 106173BR. For Basel, see Job Code 105560BR.

NIBR Presidential Postdoctoral Program – Postdoctoral positions available to study integral membrane proteins. The NIBR Presidential Postdoctoral Program provides fellows with a unique opportunity to conduct innovative, interdisciplinary basic research in an industry setting. Presidential Fellows have a NIBR mentor and an academic mentor, develop their projects in consultation with both mentors, and are expected to publish their work in leading journals.

One postdoctoral position in Basel, Switzerland, focused on the development and implementation of novel methods for large scale expression/purification of eukaryotic membrane proteins for assays, biophysics, and structure. Excellent knowledge in biochemistry of membrane proteins, secretory pathway biology, B lymphocyte biology, or lipid metabolism is desirable.

Three postdoctoral positions (1 in Basel, Switzerland and 2 in Cambridge, MA, USA) in membrane protein structural biology to further our mechanistic understanding of pharmaceutically-relevant human integral membrane protein targets. Experience in at least two of the following areas is essential: protein expression, protein purification, biophysical characterization of protein-ligand complexes, and protein structure determination. Experience working with membrane proteins is highly preferred.

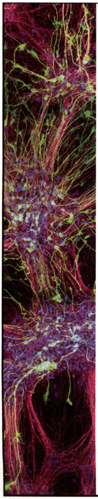
Strong collaboration skills are essential. PhD students in the last year of their doctoral research, as well as postdoctoral fellows with fewer than two years of research experience past their PhD are eligible to apply. Postdoctoral terms are for up to four years.

For more information and to apply, please visit **http://nibr.com/**. In your cover letter, reference this posting by specifying Cambridge or Basel "CPC Membrane Protein Postdoc". Candidates will be considered either for the lab head or postdoctoral positions; applications will not be considered for both types of positions.

NOVARTIS
INSTITUTES FOR
BIOMEDICAL RESEARCH

Novartis is an equal opportunity employer committed to embracing and leveraging diverse backgrounds. M/F/D/V.

NW229774R



Group Leader(s)

The Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge

Salary to £52,706 (plus start-up package)

The Gurdon Institute is a world-renowned centre in the fields of developmental, cell, and cancer biology, located in the heart of the historic city of Cambridge, and part of the University's School of Biological Sciences. Founded in 1991, its purpose is to provide the best possible environment for research, and to foster interactions and collaborations between scientists with diverse but complementary interests. It is generously supported by core funding from the Wellcome Trust and Cancer Research UK, and benefits from state-of-the art facilities in a friendly, modern, purpose-built environment (see www.gurdon.cam.ac.uk).

We are now seeking to recruit one or more new Group Leaders and are particularly interested in applicants with fewer than seven years postdoctoral experience. This would be an ideal position for a scientist wishing to take up his or her first position as an independent researcher, and offers a world-class platform for career development. Informal enquiries are invited and can be directed to any of the Institute's Group Leaders.

Applicants should send a full Curriculum Vitae and a brief outline of research interests as a single PDF file, as well as a University application form (CHRIS/6 - available from http://www.admin.cam.ac.uk/offices/hr/forms/chris6/) to glsearch@gurdon.cam.ac.uk or to the Administration Secretary at the Wellcome Trust/Cancer Research UK Gurdon Institute, Tennis Court Road, Cambridge CB2 1QN, UK, Tel: +44 1223 334133, Fax: +44 1223 334134. Applicants should also arrange for three letters of reference to be sent to the same email or mailing address. Closing date for applications: 15th December 2012.

UNIVERSITY OF CAMBRIDGE

Gurdon INSTITUTE

U230103R

JÜLICH
FORSCHUNGSZENTRUM

With a staff of 5000, Forschungszentrum Jülich – a member of the Helmholtz Association – is one of the largest interdisciplinary research centres in Europe and stands for the next generation of key technologies. Work with us on the grand challenges in the fields of health, energy & environment, and information technology, as well as on the many and varied tasks of research management.

The Institute of Bio- and Geosciences – Agrosphere (IBG-3) conducts research to improve our understanding of biogeochemical and hydrological processes in terrestrial systems. Specific studies focus on environmental controls on biogeochemical cycling of elements, the analysis of exchange processes and nutrient dynamics in the soil-plant-atmosphere continuum. A combination of experiments, modelling and innovative observation technologies is used to bridge the gap between model, process and management scale. Its research contributes to the sustainable and resource-conserving use of soils and water and to the quantification of the effect of climate and land use change on terrestrial ecosystems. We offer a competent and interdisciplinary working environment, as well as an excellent framework in the areas of experiments and modelling.

For our Institute of Bio- and Geosciences - Agrosphere (IBG-3), we are seeking at the earliest possible date for a

POSTDOC (f/m)

with a degree in natural, engineering or physical sciences

Your Job:
The overall task will be to further develop and implement analytical mass spectrometry techniques (such as MC-ICP-MS) of the use of K, Ca, Mg, Si and related isotopes and their application to enable identifying environmental controls on biogeochemical cycling of such elements. The research will contribute to provide information on the origin, fate and residence time of key elements in the terrestrial soil-plant and surface-groundwater systems, within the context of land use and climate change issues. It is envisaged that the research will make synergistic use of the existing Tereno (http://teodoor.icg.kfa-juelich.de/overview-en?set_language=en) and Soil-Can experimental facilities and infrastructure at the catchment, field and lysimeters scale.

Your Profile:

- University degree with a PhD in natural, engineering or physical sciences
- Interest in or proven experience (e.g. publications) in the use of analytical mass spectrometry techniques in environmental studies of K, Ca, Si or related isotopes
- Experience in the acquisition of funding for third-party projects as well as establishing international contacts and cooperation in the proposed field of research

We Offer:

- Competent and interdisciplinary working environment as well as excellent facilities for laboratory and field based experimental studies complimented with state of the art analytical facilities
- Employment for a fixed term of three years, with possible long-term prospects in the form of a prolongation or tenure option
- Opportunity of job-sharing
- Salary and social benefits in conformity with the provisions of the Collective Agreement for the Civil Service (TVöD)

Forschungszentrum Jülich aims to employ more women in this area and therefore particularly welcomes applications from women.

We also welcome applications from disabled persons.

Please send your written application quoting reference **233/2012NA** to:

Forschungszentrum Jülich GmbH
Geschäftsbereich Personal
- Personalentwicklung -
52425 Jülich

contact:
Please call our applicant hotline.
Phone: +49 2461 61-9700
www.fz-juelich.de

Zertifikat seit 2010
audit berufundfamilie

W230078R

Visit
www.naturejobs.com
to seriously improve your career prospects.
naturejobs
making science work

Keio University

1858
CALAMVS GLADIO FORTIOR

Tenure-Track Assistant Professor Position

at the Department of Systems Medicine, Sakaguchi Laboratory, Keio University School of Medicine, Japan

Outline: Keio University invites applicants for Assistant Professor position in the following research areas; Computational modeling and simulation of biological systems, Bioinformatics, Large-scale systematic approaches to mammalian biology.

Qualifications: Applicants must have received a Ph.D. degree within 10 years before the deadline for application, and should have expertise in the research areas listed above.

Appointment: Start of appointment: February 1, 2013 or the earliest possible date thereafter. Duration of appointment: From the date of appointment up to March 31, 2017. (Two years after the starting date, interim evaluation of research performance will be conducted. Whether employment and research support should continue will be decided based on evaluation results.)

Application Deadline: November 30, 2012, 17:00 (Japan time)

For further information and application form visit the following website:
http://www.careerpath-prj.keio.ac.jp/sakaguchi/english/
Inquiry to: keio-med-sakaguchi@adst.keio.ac.jp

JP229885R

UNIVERSITY OF NOTRE DAME
College of Science

Tenure-Track Faculty Positions in Stem Cell Biology and Rare Diseases

Stem Cells. We invite applications for three endowed junior positions in stem and iPS cell biology with interests consistent with one or more broad areas in our initiative in adult stem cell and iPS cell biology, including, but not limited to: developmental and regenerative biology, function and regulation of the stem cell niche, mechanisms underlying cellular dedifferentiation and reprogramming, regulation of gene expression in stem cells and reprogrammed iPS cells, tissue engineering, and stem cells in tumor formation and progression.
Apply by November 15.

Rare Diseases. We invite applications for an endowed junior position in rare diseases. We seek outstanding faculty candidates who use integrative and innovative approaches in the study of any rare disease, but we particularly encourage applications in areas of immune dysfunction and neurological diseases.
Apply by December 15.

science.nd.edu/faculty/jobs

The University of Notre Dame, an international Catholic research university, is an equal opportunity employer.

NW229720R

universität wien

The University of Vienna, one of Europe's oldest universities, offers approximately 9,400 employees manifold opportunities in research, teaching and administration. As stated in its development plan, the University of Vienna aims at strengthening its position as a major research-oriented university. Key elements of this strategy include the provision of an attractive range of research-based study programmes, support for furthering the work of junior academic colleagues and high-calibre professorial appointments.

The University of Vienna invites applications for the following positions.
Full details can be found at http://jobcenter.univie.ac.at/en/home/:

Tenure Track Positions

Faculty of Life Sciences

Tenure Track Position for Didactics of Biology

Tenure Track Position for Evolutionary Demography

Tenure Track Position for Plant-Microbe-Interaction

Centre for Molecular Biology

Tenure Track Position for Computational Biology
(Population Genomics/Statistical Genetics)

Closing date for applications is January 04, 2013.

If the agreed upon qualification on tenure track positions is reached, employment will be made permanent, which is connected with the title "Associate Professor" (Classification according to the Collective Bargaining Agreement: § 48 VwGr. B1 lit. b (postdoc))

The University of Vienna intends to increase the number of women on its faculty, particularly in high-level positions, and therefore specifically invites applications by women. Among equally qualified applicants women will receive preferential consideration.

W230249R

UCLA

Chief of the Division of Rheumatology at the David Geffen School of Medicine at UCLA

The Department of Medicine of the University of California at Los Angeles is seeking an exceptional leader with the expertise and skills needed to be the Chief of Rheumatology at UCLA. We are seeking a world class scientist that will continue the distinguished research that has been the hallmark of the Division over many decades. The Division includes 31 faculty members and an ACGME accredited fellowship training program consisting of 8 fellowship positions.

Individuals interested in this position should have a national/international reputation in the field of academic Rheumatology. The candidates must have a strong background in basic research in Rheumatology-related research themes with emphasis on novel approaches, a track record of extramural research funding, and publications consistent with this position. The selected individual will facilitate and coordinate all Rheumatology clinical and research activities of the Division. The candidate should have a willingness to foster collaborative interactions between clinical and laboratory scientists, and a strong interest in mentoring of junior faculty/fellows.

A tenured faculty appointment at a rank commensurate with experience will be considered.

This position will control substantial resources from the Department of Medicine, and commands a competitive salary enhanced by an attractive benefits package, including medical malpractice coverage and a collegial work environment.

The David Geffen School of Medicine at UCLA is an Affirmative Action/Equal Opportunity Employer. Women and Minorities are encouraged to apply.

Candidates should email their curriculum vitae, a letter with a statement of career goals, and the names of three references to the Chair of the Search Committee: Charalabos Pothoulakis, M.D. Professor of Medicine, David Geffen School of Medicine at UCLA, c/o Juan Vaquerano, Administrator e-mail: jvaquerano@mednet.ucla.edu

NW229876R

MRC | Laboratory of Molecular Biology

Cambridge, UK

Programme Leaders in Neurobiology

Coinciding with the move of the MRC Laboratory of Molecular Biology (LMB) into a new state-of-the-art research building from January 2013 (http://www2.mrc-lmb.cam.ac.uk/about-lmb/new-building), the Division of Neurobiology wishes to recruit talented scientists interested in developing an independent programme of research. We are particularly interested in individuals working in the area of molecular/cellular neurobiology or neurodegeneration. Current interests in the Division include mechanisms of synaptic transmission, protein misfolding in relation to neurodegenerative diseases, visual and olfactory processing and circadian rhythms (www2.mrc-lmb.cam.ac.uk/NB/). Scientific excellence and potential for major impact are paramount. Synergy with existing research programmes would be an advantage.

The LMB provides an excellent environment for hands-on research. Core-funding by the Medical Research Council provides long-term support for ambitious projects, while administrative duties are minimal and no teaching is required. Interactions across the four Divisions of the Laboratory are encouraged. There is extensive central support, including electronic and instrumentation workshops, imaging and transgenic mouse facilities. The University of Cambridge and affiliated institutions form a vibrant Neuroscience community (http://www.neuroscience.cam.ac.uk/).

Candidates should have a PhD and/or MD and will have completed a period of postdoctoral training or equivalent, with an excellent track record, and show outstanding potential for independent research. The successful candidate will lead a small team and substantial funding will be available.

These appointments will be made at either Programme Leader-track or Programme Leader level, depending on achievements and experience. Programme Leader-track appointments will be made for those who demonstrate the potential to develop into Programme Leaders within six years. Candidates for Programme Leader positions should have a strong international track record of relevant independent research and a proven ability to lead a research team, pursuing original approaches to long-term research goals. Salaries are competitive (Programme Leader-track £35,935-£48,000 per annum; Programme Leader £45,213-£70,000 per annum).

These positions will remain open until they are filled; applications received by 30th November 2012 will be given priority.

Informal enquiries can be addressed to Michel Goedert (mg@mrc-lmb.cam.ac.uk).

Applications should include a covering letter and full CV, an outline of current research interests (1 page) and a proposal for future research (up to 2 pages), along with the names and addresses of three professional referees who have agreed to be contacted prior to interview. Applications are handled by the RCUK Shared Services Centre; to apply please visit our job board at http://www.topcareer.jobs/Vacancy/irc69837_2430.aspx. If you are unable to apply online please contact us on 01793 867003 quoting reference IRC69837.

This position is subject to pre-employment screening.

For further information about the MRC visit www.mrc.ac.uk

The Medical Research Council is an Equal Opportunities Employer

U229988R

CECAD COLOGNE
Excellent in Aging Research – Exzellent in Alternsforschung

Cologne Excellence Cluster: Cellular Stress Responses in Aging-Associated Diseases announces three independent Junior Research Groups.

The mission of the Cologne Excellence Cluster on Cellular Stress Responses in Aging-associated Diseases (CECAD) will be to unravel the common molecular mechanisms underlying lifespan regulation and aging-associated diseases to set the ground to achieve its long-term objective of developing novel therapeutic interventions. CECAD is a unique research venture unifying researchers at the University of Cologne with those at the Max Planck Institutes for the Biology of Aging and Neurological Research and the German Center for Neurodegenerative Diseases (DZNE). CECAD started in 2007 and just received an extension of funding within the German Excellence Initiative. It will be active in six core research areas with a special emphasis on understanding how pathways regulating lifespan interconnect.

To further strengthen and complement the joint research program of the University of Cologne and the other partners, the Cologne Excellence Cluster CECAD announces 3 independent Junior Research Groups which are financed for five years. The position includes an attractive start-up package and access to a suite of research facilities. Candidates should have a strong track record on subjects related but not limited to mitochondrial function, genome stability, or intracellular pattern recognition in aging-associated diseases.

Applicants must include a letter of interest, a curriculum vitae, including publication list, a statement of research interests, and contact information for at least three references. Applications may be sent to the Dean of the Medical Faculty of the University of Cologne, Kerpener Str. 62, D-50931 Cologne, Germany. The closing date for applications is 22/11/2012.

We place strong emphasis on gender equality and seek to increase the proportional representation of women in this field. Thus applications from female scientists are welcome; suitably qualified women will be given preferential consideration unless other candidates clearly demonstrate superior qualifications. We also welcome applications from disabled candidates, who will also be given preferential consideration over other applicants with comparable qualifications. The position is suitable for staffing with part-time employees.

For detailed information, please visit www.cecad.uni-koeln.de

Universität zu Köln

W230242R

"*Naturejobs* has been a great way to find the right people to build our research team. In less than 1 month, I received over 40 inquiries for my advertised postdoctoral position, with many promising candidates. As a result, I have hired two very well-qualified fellows and am excited about moving our research forward with our new team."

Dr Kathleen Martin,
Yale School of Medicine, USA

naturejobs

You don't have to look this hard
to find a job with Naturejobs.

naturejobs
making science work

UMEÅ UNIVERSITET

MIMS
Molecular Infection Medicine Sweden
Nordic EMBL Partnership for Molecular Medicine

EMBL

Research Group Leaders – The Laboratory for Molecular Infection Medicine Sweden (MIMS)

The Laboratory for Molecular Infection Medicine Sweden (MIMS) was established in 2007 and is the Swedish node within the Nordic EMBL (European Molecular Biology Laboratory) Partnership for Molecular Medicine. The partnership between EMBL, the Institute for Molecular Medicine Finland (FIMM), the Centre for Molecular Medicine Norway (NCMM) and MIMS (www.mims.umu.se) is dedicated to the growing field of life sciences that investigates the molecular basis of disease and explores molecularly and genetically based treatments. The Partnership provides access to scientific infrastructure, including databases, facilities and instrumentation, as well as to clinical materials and networks and training activities provided by the partners and adopts the EMBL model for international recruitment of young group leaders, staff turnover and scientific reviews.

MIMS is established within the Umeå Centre for Microbial Research (UCMR, www.ucmr.umu.se) a highly interactive environment that includes researchers active in the fields of molecular and clinical microbiology, molecular biology, chemistry and physics and is closely connected to the university hospital.

MIMS is now seeking outstanding candidates for up to two additional Group Leader positions to be filled during 2013. The MIMS support will normally be for a five year appointment that may be extended up to a total of nine years depending on an evaluation during the fourth year. It includes a generous support package with funding for postdoctoral associates and PhD students. A successful candidate should have a record of outstanding research in fields relevant to molecular infection medicine and is expected to complement the research budgets through national and international grants. The candidate is expected to initiate and maintain strong research programs in synergy with the UCMR research environment and to take active part in collaborative research opportunities and exchange programs within the Nordic EMBL Partnership for Molecular Medicine.

More information about the EMBL Partnership and MIMS is found on www.embl.de/research/partnerships or www.mims.umu.se

We welcome your application to arrive November 25, 2012 at the latest.

Visit **www.mims.umu.se/groupleaders** for application details!

W230200R

TENURE-TRACK JUNIOR FACULTY POSITION IN SYSTEMS NEUROSCIENCE

Center for Neural Science

ARTS AND SCIENCE

The Center for Neural Science at New York University invites applications for a junior-level tenure-track faculty position from candidates with research programs in systems neuroscience. We are particularly interested in scholars whose research involves computational and/or theoretical approaches. Applications from qualified women and minorities are especially welcome. Candidates must have completed a Ph.D. or equivalent by the time of application.

The successful candidate will be expected to develop an independent extramurally funded research program, and to teach and otherwise contribute to the educational mission of the Center. The appointment can begin as early as September 1, 2013, pending administrative and budgetary approval. *The deadline for applications is* ***January 7, 2013***. To apply, see the NYU Center for Neural Science web site at **http://www.cns.nyu.edu/faculty-search**.

NEW YORK UNIVERSITY

NYU is an Equal Opportunity/Affirmative Action Employer.

NW230256R

---

nature

Chief News Editor, London

Nature, the international weekly journal of science, seeks a Chief News Editor to head its global news operation. With 15 staff reporters and a network of freelance correspondents, *Nature* produces daily online coverage and weekly print stories about all fields of research, the scientific community, policy, funding, and science and society.

The Chief News Editor plans coverage, commissions and edits stories, and works with art, production, and other *Nature* editors to deliver unique, high-impact journalism aimed at an audience of working scientists.

We seek a science journalist with extensive reporting and editing experience and a demonstrated interest in science policy and funding as well as in research findings. Management skills are essential, as is the ability to thrive under pressure.

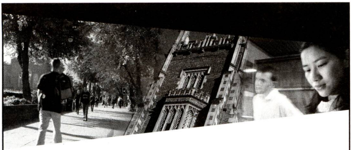
This position is offered on a full-time, permanent basis and is based in our London office.

Applicants should submit a CV and a covering letter explaining their interest in the post to vacancy **NPG/112/12** on http://jobs.macmillan.com.

Closing Date: 5th November 2012

nature publishing group npg

IN229584R



Academic Opportunities in the School of Biological Sciences

Ref: 12/102334

As part of its continuing dynamic growth, the School invites applications at all academic levels Professor, Reader, Senior Lecturer and Lecturer to complement the following areas of research: plant ecology or soil biology including plant-animal interactions and/or plant biochemistry (or cognate subject area).

For informal enquiries, please contact Professor Mark Emmerson, telephone: +44 (0) 28 9097 2912 or email: m.emmerson@qub.ac.uk.

Salary scales:

Lecturer: £32,901 - £48,246 per annum (including contribution points)
Reader/Senior Lecturer: £46,846- £59,303 per annum (including contribution points)
Professor: Salary will be in accordance with Professorial salary scales and a substantial support pack will be available

Closing date: Monday 26 November 2012

Please visit our website for further information and to apply online - **www.qub.ac.uk/jobs** or alternatively contact the address below.

The University is committed to equality of opportunity and to selection on merit. It therefore welcomes applications from all sections of society and particularly welcomes applications from people with a disability.

Personnel Department
Queen's University Belfast

Belfast, BT7 1NN.
Tel (028) 90973044
Fax (028) 90971040
E-mail on personnel@qub.ac.uk

Queen's University Belfast

RUSSELL INTERNATIONAL EXCELLENCE GROUP

A Russell Group university: one of the UK's leading research-intensive universities.

U230286R

Durham University

Shaped by the past, creating the future

BIOPHYSICAL SCIENCES INSTITUTE

RESEARCH ASSOCIATE (GRADE 6)

Salary: £23,811 - £26,779 pa **Ref: 2167**

POSTDOCTORAL RESEARCH ASSISTANT (GRADE 7)

Salary: £30,122 - £35,938 pa **Ref: 2170**

POSTDOCTORAL RESEARCH ASSOCIATE (GRADE 7)

Salary: £30,122 - £31,948 pa **Ref: 2176**

Three positions are available within the Lens Group of the Biophysical Sciences Institute at Durham University. The posts are split between two projects that will investigate eye lens structure and function. **Posts 2167 and 2170** are 3 year appointments and will combine molecular cell biology with advanced fluorescence light microscope using structured plane and multi photon application in conjunction with image segmentation approaches to model the zebrafish eye lens.

We envisage appointing individuals with biological and physical sciences backgrounds to join an interdisciplinary research team under the guidance of Professor Roy Quinlan (Biology), Professor John Girkin (Physics) and Dr Junji Wu (Engineering) working on lens regeneration, lens capsule biomechanics and imaging/segmentation/modelling tools to advance this investigation.

Post 2176 is a 1 year appointment and will address the low dose radiation effects on the lens epithelium; a candidate with experience in lens and/or radiation biology is anticipated. This appointment will be under the guidance of Professor Roy Quinlan (Biology) and Dr Liz Ainsbury (HPA Didcot).

Please contact either Professor Roy Quinlan (0191 334 1331) or Professor John Girkin (0191 334 3589) for more details. Applications from candidates interested in applying for independent fellowships are particularly encouraged.

Closing date for all posts: 22 November 2012

Further details of the posts are available on our website (http://www.dur.ac.uk/jobs/) or fax 0191 334 6495.

U230291R

---

naturejobs

Like to switch to a non-traditional science career?

- Wide selection of vacancies, such as science communications, marketing, sales, and regulatory affairs
- Browse Non-traditional career articles and news

Find out more on naturejobs.com

The Dreyfus Prize in the Chemical Sciences

The Dreyfus Prize in the Chemical Sciences recognizes an individual for exceptional and original research in a selected area of chemistry that has advanced the field in a major way. The Prize is open to individuals, both domestically and internationally.

The 2013 Dreyfus Prize will be awarded in the field of:

chemical instrumentation

The Dreyfus Prize, awarded biennially, consists of a citation, a medal, and a monetary award of $250,000.

The nomination deadline is March 1, 2013. For further information, see www.dreyfus.org.

The Camille and Henry Dreyfus Foundation, Inc.
555 Madison Avenue, 20th Floor, New York, New York 10022-3301

NW230007R

McGill Faculty of Engineering

Assistant Professor (Tenure-Track), Bioengineering

The Department of Bioengineering at McGill University invites applications for four tenure-track faculty positions. McGill University is a leading research-intensive academic institution in Canada, attracting over one-half billion dollars in competitive research funding each year. The positions are preferably at the Assistant Professor level, but outstanding senior candidates are also encouraged to apply. Rank will be determined by the qualifications of the successful applicant and salary will be commensurate with rank and experience. Successful candidates will play leading roles in establishing the Department of Bioengineering within the Faculty of Engineering and maintaining McGill's international reputation of excellence in research and teaching.

We are seeking highly qualified candidates with expertise in one or more of the following areas: biomedical microdevices, such as microarrays, lab-on-a-chip and biosensors; biomaterials, cell and tissue engineering; biological engineering, in particular concerning the applications of nanotechnology; and applications of computational and system theory to biology, biochemistry or biophysics. Candidates must have a Ph.D., preferably with a first degree in engineering, and a strong commitment to excellence in research and teaching. Evidence of outstanding research achievement is indispensable. All appointees will be expected to be, or become, members of a Canadian professional engineering association.

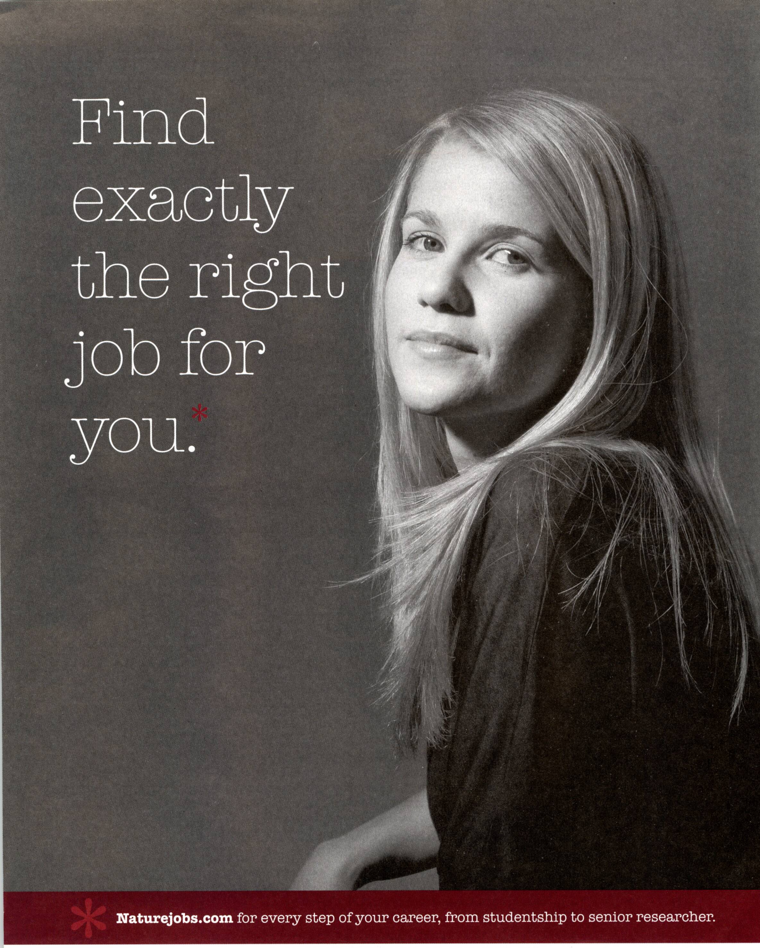
Applications will be reviewed as they are received. To ensure optimal consideration, applications must be submitted by January 31, 2013. The Faculty intends to fill the positions by August 1, 2013.

Qualified applicants are invited to submit a resume, together with names and contact information (mail, phone and email) of three references, a two-page statement outlining research and teaching goals, and copies of recent publications to: Professor Dan V. Nicolau, Chair, Department of Bioengineering, McGill University, 817 Sherbrooke Street West, Room 378, Montreal, Quebec, Canada H3A 0C3 or by e-mail to facultysearch.bioeng@mcgill.ca

All qualified applicants are encouraged to apply; however, Canadians and permanent residents will be given priority. McGill University is committed to equity in employment and diversity. It welcomes applications from indigenous peoples, visible minorities, ethnic minorities, persons with disabilities, women, persons of minority sexual orientations and gender identities and others who may contribute to further diversification.

Please reference the source of the ad when applying for, or inquiring about, this job announcement.

NW229781R


Find
exactly
the right
job for
you.*

* **Naturejobs.com** for every step of your career, from studentship to senior researcher.

naturejobs

nature publishing group npg

Need to find the ideal candidate *fast?*

Visit **www.naturejobs.com** to discover how applicants can respond directly to you by email.

naturejobs
making science work

FACULTY POSITIONS

in the College of Chemistry and Molecular Engineering at Peking University

The College of Chemistry and Molecular Engineering (CCME) at Peking University invites applications for faculty positions, with preference for tenure-track positions for juniors through the"Thousand Young Talents Program" initiated and sponsored by the Chinese government. We particularly welcome creative and energetic individuals with strong backgrounds in all subdisciplines of chemistry.

CCME, formerly known as the Department of Chemistry of Peking University, was the first chemistry department founded in China in 1910. It embraces both conventional and interdisciplinary areas of chemistry, including inorganic, organic, analytical, physical, theoretical and computational chemistry, materials chemistry, polymer science and engineering, applied chemistry, and chemical biology. For detailed information about CCME, please visit our website at **http://www.chem.pku.edu.cn.**

Applicants should have a strong research profile to establish and conduct innovative research in chemistry or relating fields. CCME provides state-of-the-art research facilities and supporting system. Internationally competitive start-up packages, salary and benefits will be offered according to qualifications and experiene.

Application materials should include: 1) a curriculum vitae; 2) a five-year research plan; 3) three references with detailed contact information; 4) three to five representative publications in PDF format. All materials should be sent via e-mail to **ghxhr@pku.edu.cn** or regular mails to: Human Resource Office, Room C111, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China.

JP223118R

Memorial Sloan-Kettering Cancer Center
The Best Cancer Care. Anywhere.

PostDoc - Quantitative Proteomics

Join Ronald Hendrickson's newly formed research group as a PostDoc in Molecular Pharmacology and Chemistry at Memorial Sloan-Kettering Cancer Center (MSKCC) and Sloan Kettering Institute (SKI) in New York City. Co-located with the Proteomics Core facility in the Rockefeller Research Building and adjacent to Memorial Hospital and the Zuckerman Research Center at MSKCC, this rich scientific and collaborative environment fosters innovation to solve critical problems in cancer biology and to discover more effective treatments for cancer.

Our dynamic research team is focused on developing and applying state-of-the-art proteomic methods at the intersection where basic science and clinical research meet.

Current and future research & development topics include:

- mass spectrometric measurement of activated PI3K pathways and feedback mechanisms in cancer cell lines and core needle tumor biopsies
- protein flux and dynamics
- top down analysis of protein isoforms and PTMs

Join a newly formed and highly motivated group of biochemists and mass spectrometrists. Collaborate with others and lead a project. Do imaginative basic science. Apply your results to disease biology and health care. Candidates must have a Ph.D. or equivalent in biology, biochemistry, medicine or mathematics. Please send CV, bibliography, brief statement of research interests and names of 3 references to **donnarud@mskcc.org**.

facultysearch.ski.edu

MSKCC is an equal opportunity and affirmative action employer committed to diversity and inclusion in all aspects of recruiting and employment. All qualified individuals are encouraged to apply.

NW230322R

2013 HEINRICH WIELAND PRIZE

HEINRICH WIELAND 1877-1957

The Boehringer Ingelheim Foundation invites nominations for the international

HEINRICH WIELAND PRIZE 2013.

The

50,000 Euro

prize is awarded annually by an independent board of trustees to honour and recognize outstanding research on biologically active molecules and systems in the areas of chemistry, biochemistry and physiology as well as on their clinical impact. The prize is named after Heinrich Wieland, a Nobel Prize Laureate in chemistry, and presented every autumn in Munich, Germany. (www.heinrich-wieland-prize.de)

NOMINATIONS MUST BE SUBMITTED BY FEBRUARY 28, 2013

and include the following documents (in English):
1) in hard copy: laudation of no more than two pages summarizing the achievements of the candidate, CV and list of the five most relevant publications of the past 10 years,
2) on CD: bibliography and manuscripts of the five listed publications.

Please send the documents to Dr. Anja Hoffmann, Boehringer Ingelheim Foundation, Schusterstr. 46-48, 55116 Mainz, Germany.

The Boehringer Ingelheim Foundation (www.boehringer-ingelheim-stiftung.de) is an independent, non-profit institution for the promotion of research in medicine, biology, chemistry and pharmacy.

Boehringer Ingelheim Stiftung

W230282AM

CORFO
sueña emprende crece

Creative BioScience

>>> INVITATION TO TENDER

Uxmal S.A. is seeking companies or institutional organizations interested in

- Performing the chemical characterization for an isolated peak of a the new molecule using. NMR (TOCSY, COSY HSQC, 1H y 13C, HMBC 1H, 13C, NOESY &/o ROESY).
  Goal: Chemical structure
- Performing the chemical structural synthesis.
  Goal: 3 grams synthetic new compound.

CORFO INNOVA 12IDL2-13648

Your application should include costs, details of services to be provided and experience as well as any other details that you consider relevant in the decision making process to be undertaken by Uxmal's experts.
Closing Date: November 15th, 2012
Institutions or companies interested to applying please send your proposal as a single PDF to Manuel Gidekel, General Manager of Uxmal S.A. Chile, at the following email:
mgidekel@creativebio-science.com
Informal enquiries should also be addressed to Dr. Gidekel.

Note:

The institutions or companies interested in:

- performing to discover the novel structure and synthesis of the new molecule. The company or institute should have the capacity and be authorized by regulatory bodies, either from Europe or the US, to produce the chemical batches.
- **NDA are required**

Applicants can select one or two areas of the call.

CORFO is the National Development Agency reporting directly to the Chilean Government's Ministry of Economy. Uxmal S.A. is a Chilean biotech start up aiming to discover and develop novel therapeutic compounds for cancer. Uxmal S.A. is part of the Creative BioScience Platform, supported by private funds & grants from the Chilean Government.

NW230041A

Don't miss the intoxicatingly good job opportunities in *Nature* each week and on naturejobs.com

naturejobs

INVESTISSEMENTS D'AVENIR

Institut Pasteur

Call for Candidates to a chair in Integrative Biology of Infection

The Institut Pasteur announces an international call for a Chair of "Infection Systems Biology" on its Paris campus. The position is open for an outstanding candidate who wishes to develop a cutting-edge research program on microbial pathogenesis (caused by virus, bacteria, fungi or protozoan parasites) with a strong focus/interest on systems-level proteomic approaches on any aspect of infection biology.

The chair holder will join the Laboratory of Excellence (LabEx) program called, "Integrative Biology of Emerging Infectious Diseases" (IBEID: http://www.pasteur.fr/labex/ibeid) and will benefit from the cutting edge interdisciplinary environment provided by the Institut Pasteur in Paris and by the LabEx.

The Chair will also benefit from the technical and scientific expertise of the Institut Pasteur proteomics core facility. He/she will have access to state-of-the-art equipment in mass spectrometry (last generation orbitraps equipped with ETD, MALDI-TOF/TOF, triple quadrupole) and separation techniques (nanoLC, UPLC, 2D gel electrophoresis) allowing high throughput data to be obtained. The platform has been completely renovated and is now part of a research unit in structural mass spectrometry and proteomics, with which interactions will be facilitated.

The candidate should have a PhD and/or MD degree, relevant experience as an independent investigator and a demonstrated track record of high productivity.

The deadline for applications is April 1, 2013. Short-listed candidates will be called for interview in May-June 2013, and decisions will be announced by October 2013.

General conditions and applications. The successful candidate will be appointed as Chair of Infection Systems Biology for a period of 8 years and will create an Infection Systems Biology research unit. The budget (up to 3,200,000€ over 8 years) includes a tenured salary for the Chair (if necessary), 3 three-year postdoctoral positions, a technician, half-time secretarial assistance, basic laboratory equipment, a substantial contribution to running costs and essential large equipment, and access to on-campus facilities including state-of-the-art technology platforms.

Candidates should send their formal applications by E-mail to the Director of Scientific Evaluation, Prof. Alain Israël, at the Institut Pasteur (25, rue du Dr. Roux, 75724 Paris, France; ibeidchair@pasteur.fr).

The Application shall comprise the following (in order) in a single pdf file:

1. A brief introductory letter (candidates are encouraged to contact the IBEID coordinators, Pascale Cossart (pascale.cossart@pasteur.fr) or Philippe Sansonetti (philippe.sansonetti@pasteur.fr))
2. A Curriculum Vitae and a full publication list.
3. A description of past and present research activities (up to 5 pages with 1.5 spacing).
4. The proposed research project (up to 10 pages with 1.5 spacing) and how it would fit in the defined topic.
5. The names of 3 scientists from whom letters of recommendation can be sought, together with the names of scientists with a potential conflict of interest from whom opinions should not be requested.

W230147A

Achieving food and environmental security

New approaches to close the gap

3 – 4 December 2012
The Royal Society, 6 – 9 Carlton House Terrace,
London, SW1Y 5AG

Organised by Professor Guy Poppy, Professor Paul Jepson, Professor John Pickett CBE FRS and Dr Michael Birkett

A growing population coupled with climate change increases pressure on land and food security. The sustainable production of sufficient food, water and energy is challenging and boundaries between intensive/extensive and high tech/low tech agriculture need to be broken down. This meeting reports on innovative science and adoption of novel strategies that open new windows of opportunity for sustainable production intensification.

This event is free to attend. Places are limited and pre-registration online is essential. For more information, visit **royalsociety.org/events/2012/close-the-gap**

THE
ROYAL
SOCIETY

Registered Charity No 207043

U229854EL

Visit **www.naturejobs.com** to seriously improve your career prospects.

naturejobs
making science work

REPROGRAPHICS: CC Media Group, London EC1V 8AR UK and **The Charlesworth Group,** Wakefield, UK.
PRINTED BY: William Gibbons & Sons Ltd, UK; **Publishers Press**, Lebanon Junction, Ky, USA and **Obun Printing Co. Inc,** Tokyo, Japan

a nature conference

University of Miami, International Union of Biochemistry and Molecular Biology, Nature Publishing Group and Scripps Florida present:

Miami 2013 Winter Symposium: The Molecular Basis of Metabolism and Nutrition

Join leaders in the field to discuss advances in our molecular understanding of obesity, calorie restriction, diabetes, ageing and the microbiome as they relate to human nutrition.

February 10-13, 2013
JW Marriott Marquis Miami
Miami, FL, USA

SYMPOSIUM AWARDEES

Feodor Lynen Lectures
Ana Maria Cuervo (Albert Einstein College of Medicine, USA)
Leonard Guarente (Massachusetts Institute of Technology, USA)
David Sabatini (Whitehead Institute for Biomedical Research, USA)

Distinguished Service Award
Bruce Spiegelman (Harvard University, USA)

Special Achievement Awards
Lewis Cantley (Harvard University, USA)

IUBMB Jubilee Lecture
Michael Karin (University of California, San Diego, USA)

CONFIRMED SPEAKERS:

Arthur Agatston
Fredrik Bäckhed
Anne Brunet
Tom Burris
John Cleveland
Miguel Constância
Ronald DePinho
Richard Flavell
Stanley Hazen
Jules Hirsch
Randal Kaufman
Salvador Moncada
Deborah Muoio
Chris Newgard
Jared Rutter
Reuben Shaw
Matthew Vander Heiden
Douglas Wallace
Robert Waterland

For more information and to register visit:
www.nature.com/natureconferences/miami/mws2013

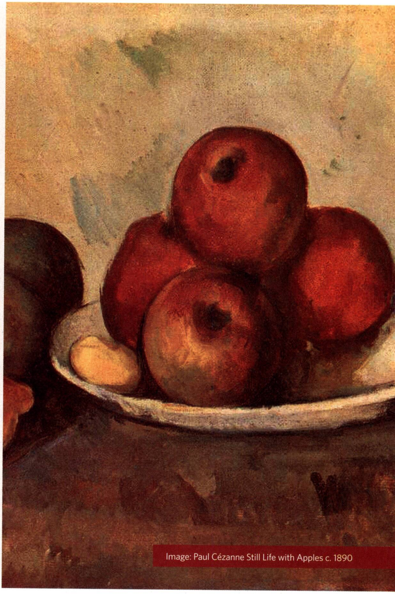


Image: Paul Cézanne Still Life with Apples c. 1890

UNIVERSITY OF MIAMI
MILLER SCHOOL
of MEDICINE

SCRIPPS
FLORIDA™
THE SCRIPPS RESEARCH INSTITUTE

IUBMB

IN228654E

MIDNIGHT IN THE CATHEDRAL OF TIME

Dream sequence.

BY PRESTON GRASSMANN

Walking through the crowded streets in Canvas Town, I pass a booth that claims to sell the god-ware of angelic systems; codes that open gateways to palaces of corporate data. I pass the aisles of snake-oil salesmen, hawking the latest nanotech cures for assorted ailments, from back pain to cancer. Through the aisles, the gold and silver relics of archaic religions are nestled among the silicon and plastic wafers of data-chips.

The man I'm searching for sits in the corner of the tent and looks up slowly. His eyes catch a glint of light from mechanically modified plasma-eels that swim in a tank at the entrance of his shop. As I enter, he smiles and hunches slowly forward, a conspirator waiting to whisper a secret. Towers fall and cathedrals break apart across his chest, streets narrow across his ribs to make room for buildings; his skin altering in the movement.

"You finally made it back," he says, his lips turning a mountain ravine into a cave of broken stalactites. He watches me with grey-green eyes.

"What made you think I would?" I ask, watching as two eels collide in a surge of blue light, plasma flickering around their bodies. A few people stand at the entrance, placing their hands on the surface of the tank to watch the storm surge around their fingers.

"Lucid dreamers always return."

I begin to see the illustrations spreading outwards from his body, an artist's outline that turns into solid shapes. The mountain across his cheekbones begins to melt, spreading down his chest, colours pouring across his skin like watercolours.

The DJ stays where he is, unmoving, as if waiting for me to act.

"One reading is all it takes for some," he says, reaching out for the data-chip. I place it in his open palm. "Before they become…"

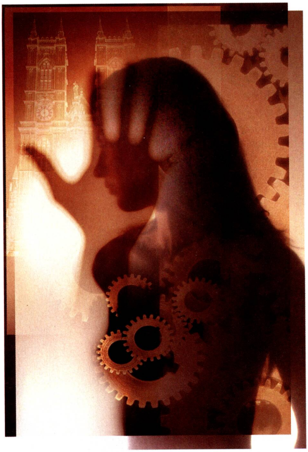
"Dream junkies?"

He nods slowly.

A crowd begins to gather outside the shop. He replays the images of dreams from previous customers. Another image forms across his chest; bright green palms swaying among a village of burning homes, spires of smoke rising across his shoulders like wings.

➲ NATURE.COM
Follow Futures:
@NatureFutures
go.nature.com/mtoodm

The data-chips were dreams rendered into code. There were only a few technicians who understood how such data could be used, and even fewer who could render them into coherent forms. This man had chosen to turn it into an art. As he read the data-card, the images would be distilled and remixed, projected on his skin in high resolution.



JACEY

"Shall we keep this private?" he asks, the pictures on his skin shifting too quickly to be seen.

"Yes."

He rises for a moment to move the audience away and pulls a low Japanese panel-screen across the entrance.

"There are more details here than before," he says, as the shadows across his chest begin to resolve, her face becoming clearer, chiaroscuro lines opening into the shadowed hollows of her eyes, the twist of a red mouth, the familiar angles of her cheeks. She stands at the entrance. "It's midnight…," she says, the voice of the DJ altering to let hers come through.

She is wraith-thin and dressed in a silk robe and she retreats back into the cathedral, moving as if part of some clockwork mechanism. Through a sequence of images, I watch myself enter, searching for her through the aisles. Everything inside is turning with rusted clockwork parts and the corroded mandalas of half-broken gears. In the pews, worshippers hold the remains of broken clocks. The springs and cogs are spread across their open palms like hymnals.

"Horo-shippers," the DJ says softly, smiling. "Clever."

And then I find her, staring up at the altar. She doesn't turn away as I approach. Even as I put my hand on her shoulder, she remains still. All at once, everything stops. The DJ is motionless, his smile fading as the images of the gears stop turning and the parts fall out of the worshippers' hands and clatter on the floor. And then there is only stillness inside and I run to the altar and beat my fists against its walls, striking out against the mechanism. "Not yet," I say, my own voice echoing around me. "Come back."

That's when the face of the man begins to shift and flow, and she is there, the slope of her cheeks, the curve of her mouth, the soft line of her eyebrows, eyes rendered in the space of eyelids; the same bright, piercing blue-green that I remember so well. And then her body forms, shaping itself across the terrain of his arms, his chest, flowing down the length of his body, until the DJ's skin is only a faint outline around hers. For a moment I can still see the gears, as if she has become the cathedral, the movements of time, her heart the mechanism that makes the gears turn, but they soon fade until it is only her body reaching out for mine.

"Come," she says, and though I know this can't be part of my recorded dream, that it's only an offering, I will hold her for as long as I can. ■

Preston Grassmann *is co-author of* The Floating World, *a collaboration with K. J. Bishop. He currently lives in Japan, surrounded by a growing collection of drawings, paintings and unreliable maps of the Tokyo underground.*

7,000 Human Proteins Made from Human Cells

More Assays!
Better Assays!

30% OFF*

The Largest Collection of Mammalian Cell-Derived Proteins

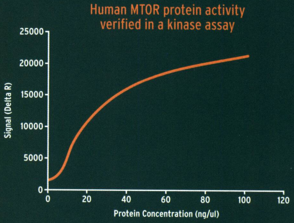
Protein functional studies

Compound screening

Immunogens for antibody production

Assay standards: ELISA, Mass Spec, and more

Bulk production available



*3000 proteins, expires 12/15/12

origene.com/protein

ORIGENE
Your Gene Company

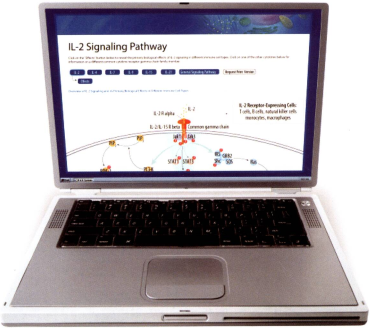
R&D Systems Tools for Cell Biology Research™

NEW Interactive Pathways & Processes

✓ Designed by scientists for scientists

✓ One click to access the widest collection of products

✓ Request a printed copy, which includes additional resources


IL-2 Signaling Pathway
IL-2
IL-4
IL-7
IL-9
IL-15
IL-21
General Signaling Pathway
Request Print Version
Effects
IL-2 Receptor-Expressing Cells:
T cells, B cells, natural killer cells
monocytes, macrophages
IL-2
IL-2 R alpha
IL-2/IL-15 R beta
Common gamma chain
Jak1
Jak3
PIP
PIP
STAT1
STAT3
IRS
GRB2
Shc
SOS
Ras
PI3-K

R&D Systems is excited to announce the addition of Interactive Pathways & Processes to the growing list of reference tools available on our website. These new fully-interactive pathways allow you to explore a variety of biological processes, including Apoptosis, Wnt Signaling, and Th1 Differentiation.

Start exploring today
www.RnDSystems.com/**Pathways**

R&D SYSTEMS®